

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16318 A2

(51) International Patent Classification⁷: C12N 15/12,
C07K 14/47, 14/705, G01N 33/53, C12N 15/62, C07K
16/18

PCT/US00/04414 22 February 2000 (22.02.2000) US
PCT/US00/05601 1 March 2000 (01.03.2000) US
60/187,202 3 March 2000 (03.03.2000) US
60/191,007 21 March 2000 (21.03.2000) US
PCT/US00/08439 30 March 2000 (30.03.2000) US
60/199,397 25 April 2000 (25.04.2000) US
PCT/US00/14042 22 May 2000 (22.05.2000) US
60/209,832 5 June 2000 (05.06.2000) US

(21) International Application Number: PCT/US00/23328

(22) International Filing Date: 24 August 2000 (24.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PCT/US99/20111

PCT/US99/21090 1 September 1999 (01.09.1999) US

60/169,495 15 September 1999 (15.09.1999) US

60/170,262 7 December 1999 (07.12.1999) US

60/175,481 9 December 1999 (09.12.1999) US

PCT/US00/04341 11 January 2000 (11.01.2000) US

PCT/US00/04342 18 February 2000 (18.02.2000) US

PCT/US00/04342 18 February 2000 (18.02.2000) US

(71) Applicant (for all designated States except US): GENEN-TECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): EATON, Dan, L. [US/US]; 75 Knight Drive, San Rafael, CA 94901 (US). FILVAROFF, Ellen [US/US]; 538 18th Avenue, San Francisco, CA 94121 (US). GERRITSEN, Mary, E. [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). GRIMALDI, Christopher, J. [US/US]; 1434 36th

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

<subunit 1 of 1, 266 aa, 1 stop
<MW: 29766, pI: 8.39, NX(S/T): 0
MWWFQQQLSFLPSALVIWTSAAFIFSYSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIA
AVLCIATIYVRYKQVHALSPEENVIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVGAV
LTFGMGSLYMFVQTILSYQMOPKIHGKQVFWRLLLVIWCGVSALSMLTCSVLHSGNFGTDL
EQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPC
PINNERTRLLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

A4 - 09/925,055

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

WO 01/16318 A2

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE
SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the
5 recombinant production of novel polypeptides.

BACKGROUND OF THE INVENTION

Extracellular proteins play important roles in, among other things, the formation, differentiation and
10 maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration,
differentiation, or interaction with other cells, is typically governed by information received from other cells
and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance,
mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which
are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted
15 polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of
action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics,
biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons,
interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins.
Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts
20 are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are
focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel
secreted proteins. Examples of screening methods and techniques are described in the literature [see, for
example, Klein et al., Proc. Natl. Acad. Sci. 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation,
25 differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation,
migration, differentiation, or interaction with other cells, is typically governed by information received from
other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides
(for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and
hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins.
30 Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor
kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like
selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is
regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze
that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and

nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein,
5 the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid
10 sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90%
15 nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively
20 at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-
25 inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide
30 that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length,
35 alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length.

amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO300 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA40625-1189".

Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

5 Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA45409-2511".

Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO1282 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA45495-1550".

10 Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO1063 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA49820-1427".

15 Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO1773 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA56406-1704".

Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

20 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO1013 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA56410-1414".

Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

25 Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO937 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA56436-1448".

Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO842 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA56855-1447".

30 Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO1180 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA56860-1510".

35 Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO831 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA56862-1343".

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO1069 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA59211-1450".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

5 Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA59212-1627".

Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

10 Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO1129 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA59213-1487".

Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO1027 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA59605-1418".

15 Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO1106 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA59609-1470".

Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

20 Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO1291 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA59610-1556".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

25 Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO3573 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA59837-2545".

Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO3566 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA59844-2542".

30 Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO1098 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA59854-1459".

35 Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO1158 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA60625-1507".

Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO1270 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA66308-1537".

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

5 Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO1268 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA66519-1535".

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

10 Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO1327 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA66521-1583".

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO1328 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA66658-1584".

15 Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA66660-1585".

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

20 Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO1340 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA66663-1598".

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

25 Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO1342 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA66674-1599".

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO3579 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA68862-2546".

30 Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO1472 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA68866-1644".

35 Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO1461 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA68871-1638".

Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO1566 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA77568-1626".

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

5 Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO1774 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA77626-1705".

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA81754-2532".

10 Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO1865 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA81757-2512".

15 Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO1925 cDNA, wherein SEQ ID NO:133 is a clone designated herein as "DNA82302-2529".

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

20 Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO1926 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA82340-2530".

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

25 Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA83500-2506".

Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO4405 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA84920-2614".

30 Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO3435 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA85066-2534".

35 Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO3543 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA86571-2551".

Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO20233 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA165608".

Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

5 Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO19670 cDNA, wherein SEQ ID NO:165 is a clone designated herein as "DNA131639-2874".

Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1890 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA79230-2525".

10 Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

15 The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino

sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which 5 may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-10 encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those 15 search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can 20 alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

25 where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

30 In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

35 "Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous

complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any

forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain (V_H-V_L). By using a linker that is too short to allow pairing between the two domains

Table 1

40

45

50

55

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 * where file1 and file2 are two dna or two protein sequences.
 * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
{
    int ac;
    char *av[ ];
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */

    nw(); /* fill in the matrix, get the possible jmps */
    readjmps(); /* get the actual jmps */
    print(); /* print stats, alignment */

    cleanup(0); /* unlink any tmp files */
}

```

Table 1 (cont')

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongoing del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongoing del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score: we're favoring
     * mis over any del and delx over dely
     */
}

55

60

```

Table 1 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 *
5   * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[ ]; print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
10  * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE   256      /* maximum output line */
#define P_SPC    3      /* space between name or num and seq */

20 extern _day[26][26];
int olen;           /* set output line length */
FILE *fx;           /* output file */

25 print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

30    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    sprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    sprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
35    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
40        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
45        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
50        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
55        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

60

```

Table 1 (cont')

```

    fprintf(fx, "< gaps in first sequence: %d", gapx); ...getmat
    if (gapx) {
        (void) sprintf(outx, " (%d %s%s)", 5
                     ngapx, (dna)? "base":"residue", (ngapx == 1)? ":";"s");
        fprintf(fx,"%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy); 10
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)", 10
                     ngapy, (dna)? "base":"residue", (ngapy == 1)? ":";"s");
        fprintf(fx,"%s", outx);
    }
    if (dna) 15
        fprintf(fx,
                 "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
                 smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx, 20
                 "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                 smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx, 25
                 " < endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                 firstgap, (dna)? "base" : "residue", (firstgap == 1)? ":";"s",
                 lastgap, (dna)? "base" : "residue", (lastgap == 1)? ":";"s");
    else
        fprintf(fx, " < endgaps not penalized\n");
}
30
static nm; /* matches in core -- for checking */
static lmax; /* lengths of stripped file names */
static ij[2]; /* jmp index for a path */
static nc[2]; /* number at start of current line */
static ni[2]; /* current elem number -- for gapping */
static siz[2];
static char *ps[2]; /* ptr to current element */
static char *po[2]; /* ptr to next output char slot */
40 static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp [ ]
 */
45 static pr_align()
{
    int nn; /* char count */
    int more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(namex[i]);
        if (nn > lmax)
            lmax = nn;
55
        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}
60
pr_align

```

Table 1 (cont')

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' '))
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
    putline(i);
    if (i == 0 && *out[1])
        fprintf(fx, star);
    if (i == 1)
        nums(i);
}
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)
int      ix;      /* index in out[] holding seq line */
{
    char      nline[P_LINE];
    register   i, j;
    register char  *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax + P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)
int      ix;
{

```

nums

putline

Table 1 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
5   stripname(pn)
      char    *pn;     /* file name (may be path) */
{
    register char    *px, *py;

10  py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
15    (void) strcpy(pn, py);
    return(strlen(pn));
}
20

25

30

35

40

45

50

55

60
```

Table 1 (cont')

```

...getseq
5      py = pseq + 4;
*len = tlen;
rewind(fp);
10     while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
            if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU", *(py-1)))
                natgc++;
        }
        *py++ = '\0';
        *py = '\0';
20     (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
    }

25     char * g_calloc(msg, nx, sz)
30     g_calloc(msg, nx, sz)
        char *msg; /* program, calling routine */
        int nx, sz; /* number and size of elements */
    {
        char *px, *calloc();
        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
            if (*msg) {
                fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                exit(1);
            }
        }
        return(px);
    }

40     /*
45     * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
46     */
readjmps()
47     {
        int fd = -1;
        int siz, i0, i1;
        register i, j, xx;

50     if (fj) {
            (void) fclose(fj);
            if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                cleanup(1);
            }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
            while (1) {
                for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                    ;
            }
        }
    }

```

readjmps

Table 1 (cont')

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5   writejmps(ix)
     int      ix;
{
    char    *mktemp();

10  if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
15  if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
20  (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

25

30

35

40

45

50

55

60

```

Table 3

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXYYYYZZYZ	(Length = 15 amino acids)

5 % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10 5 divided by 10 = 50%

1

Table 5

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10 4 divided by 12 = 33.3%

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide
5 fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired
10 termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial
15 changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

10

C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the

D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart 5 et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length 10 PRO.

1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently 15 obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA 20 or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences 25 selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like 32 P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known 30 sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic 35 libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

ilvG kan'; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning 5 or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, *Nature*, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., *Bio/Technology*, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., *J. Bacteriol.*, 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilicola* (ATCC 36,906; Van den Berg et al., *Bio/Technology*, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., *J. Basic Microbiol.*, 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., *Proc. Natl. Acad. Sci. USA*, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, 10 e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., *Biochem. Biophys. Res. Commun.*, 112:284-289 [1983]; Tilburn et al., *Gene*, 26:205-221 [1983]; Yelton et al., *Proc. Natl. Acad. Sci. USA*, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, *EMBO J.*, 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, 15 *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary 20 of this class of yeasts may be found in C. Anthony, *The Biochemistry of Methylotrophs*, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila S2* and *Spodoptera Sf9*, as well as plant 25 cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); human lung 30 cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector 35 may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally

promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess 5 et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription 10 controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters 15 obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an 20 enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication 25 origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated 30 cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., 35 Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native
5 nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the
10 probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using
15 the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14
20 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including
25 enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of
30 resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increase affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or
35 antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of

analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 µg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO

If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers
5 (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system")
10 takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-lacZ reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β-galactosidase. A complete kit
15 (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein
20 and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation)
25 between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be
30 screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide
35 molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991).

polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, *Current Biology*, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

25

1. Polyclonal Antibodies

The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂, or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones

pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in 10 at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody 15 molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large 20 amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ 25 bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB 30 derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized 35 bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., 176: 1191-1195 (1992) and Shope, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* 5 Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

7. Immunoconjugates

10 The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active 15 fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the trichothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

20 Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis-(*p*-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(*p*-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent 25 for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for 30 utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

8. Immunoliposomes

35 The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos.

acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture 5 at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulphydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

10

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, e.g., detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as 15 competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ³H, ¹⁴C, ³²P, ³⁵S, or ¹²⁵I, a fluorescent or 20 chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

25 Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized 30 antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

35 All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLE 2: Isolation of cDNA clones by Amylase Screening1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System).

5 In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkerered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

10 A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linkerered with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI 15 cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

20 3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was 25 added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the 30 plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT 35 alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in sec71, sec72, sec62, with truncated sec71 being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 µl) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 µl) was used as a template for the PCR reaction in a 25 µl volume containing: 0.5 µl Klentaq (Clontech, Palo Alto, CA); 4.0 µl 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 µl Kentaq buffer (Clontech); 0.25 µl forward oligo 1; 0.25 µl reverse oligo 2; 12.5 µl distilled water.

The sequence of the forward oligonucleotide 1 was:

5'-TGTAAAACGACGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:169)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:170)

10 PCR was then performed as follows:

a.		Denature	92°C, 5 minutes
b.	3 cycles of:	Denature	92°C, 30 seconds
		Anneal	59°C, 30 seconds
		Extend	72°C, 60 seconds
c.	3 cycles of:	Denature	92°C, 30 seconds
		Anneal	57°C, 30 seconds
		Extend	72°C, 60 seconds
d.	25 cycles of:	Denature	92°C, 30 seconds
		Anneal	55°C, 30 seconds
		Extend	72°C, 60 seconds
25	e.	Hold	4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine

Table 7 (cont')

5	DNA60625-1507	209975	June 16, 1998
	DNA60629-1481	209979	June 16, 1998
	DNA61755-1554	203112	August 11, 1998
5	DNA62812-1594	203248	September 9, 1998
	DNA62815-1576	203247	September 9, 1998
	DNA64881-1602	203240	September 9, 1998
	DNA64886-1601	203241	September 9, 1998
10	DNA64902-1667	203317	October 6, 1998
	DNA64950-1590	203224	September 15, 1998
	DNA65403-1565	203230	September 15, 1998
	DNA66308-1537	203159	August 25, 1998
	DNA66519-1535	203236	September 15, 1998
15	DNA66521-1583	203225	September 15, 1998
	DNA66658-1584	203229	September 15, 1998
	DNA66660-1585	203279	September 22, 1998
	DNA66663-1598	203268	September 22, 1998
	DNA66674-1599	203281	September 22, 1998
20	DNA68862-2546	203652	February 9, 1999
	DNA68866-1644	203283	September 22, 1998
	DNA68871-1638	203280	September 22, 1998
	DNA68880-1676	203319	October 6, 1998
	DNA68883-1691	203535	December 15, 1998
25	DNA68885-1678	203311	October 6, 1998
	DNA71277-1636	203285	September 22, 1998
	DNA73727-1673	203459	November 3, 1998
	DNA73734-1680	203363	October 20, 1998
	DNA73735-1681	203356	October 20, 1998
30	DNA76393-1664	203323	October 6, 1998
	DNA77301-1708	203407	October 27, 1998
	DNA77568-1626	203134	August 18, 1998
	DNA77626-1705	203536	December 15, 1998
	DNA81754-2532	203542	December 15, 1998
35	DNA81757-2512	203543	December 15, 1998
	DNA82302-2529	203534	December 15, 1998
	DNA82340-2530	203547	December 22, 1998
	DNA83500-2506	203391	October 29, 1998
	DNA84920-2614	203966	April 27, 1999
40	DNA85066-2534	203588	January 12, 1999
	DNA86571-2551	203660	February 9, 1999
	DNA87991-2540	203656	February 9, 1999
	DNA92238-2539	203602	January 20, 1999
	DNA96042-2682	PTA-382	July 20, 1999
45	DNA96787-2534	203589	January 12, 1999
	DNA125185-2806	PTA-1031	December 7, 1999
	DNA147531-2821	PTA-1185	January 11, 2000
	DNA115291-2681	PTA-202	June 8, 1999
	DNA164625-28890	PTA-1535	March 21, 2000
50	DNA131639-2874	PTA-1784	April 25, 2000
	DNA79230-2525	203549	December 22, 1998

These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

5 Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

10 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

15 PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq)). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco 20 yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

25 *E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The 30 clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

35 The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The

serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 78:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 µg pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 µg/ml bovine insulin and 0.1 µg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

EXAMPLE 8: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 9: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfected the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 µm

affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 11: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide 5 is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.).

10 Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

15 Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

20 A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

25 EXAMPLE 12: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located 30 intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex 35 formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with

antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available
5 to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 14: Pericyte c-Fos Induction (Assay 93)

This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Induction of c-fos expression in pericytes is also indicative of the induction of angiogenesis and, as such, PRO polypeptides capable of inducing the expression of c-fos would be expected to
10 be useful for the treatment of conditions where induced angiogenesis would be beneficial including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of PRO polypeptide
15 test samples and controls (positive control = DME+5% serum +/- PDGF at 500 ng/ml; negative control = protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control)
20 value indicated by chemiluminescence units (RLU) luminometer reading versus frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media = low glucose DMEM = 20% FBS + 1X pen strep + 1X fungizone. Assay Media = low glucose DMEM + 5% FBS.

25 The following polypeptides tested positive in this assay: PRO1347 and PRO1340.

EXAMPLE 15: Ability of PRO Polypeptides to Stimulate the Release of Proteoglycans from Cartilage (Assay 97)

The ability of various PRO polypeptides to stimulate the release of proteoglycans from cartilage tissue
30 was tested as follows.

The metacarpophalangeal joint of 4-6 month old pigs was aseptically dissected, and articular cartilage was removed by free hand slicing being careful to avoid the underlying bone. The cartilage was minced and cultured in bulk for 24 hours in a humidified atmosphere of 95% air, 5% CO₂ in serum free (SF) media (DME/F12 1:1) with 0.1% BSA and 100U/ml penicillin and 100 μ g/ml streptomycin. After washing three
35 times, approximately 100 mg of articular cartilage was aliquoted into micronics tubes and incubated for an additional 24 hours in the above SF media. PRO polypeptides were then added at 1% either alone or in combination with 18 ng/ml interleukin-1 α , a known stimulator of proteoglycan release from cartilage tissue.

The following PRO polypeptides tested positive in this assay: PRO263, PRO295, PRO1282, PRO1063, PRO1356, PRO3543, and PRO5990.

EXAMPLE 18: Tumor Versus Normal Differential Tissue Expression Distribution

Oligonucleotide probes were constructed from some of the PRO polypeptide-encoding nucleotide sequences shown in the accompanying figures for use in quantitative PCR amplification reactions. The oligonucleotide probes were chosen so as to give an approximately 200-600 base pair amplified fragment from the 3' end of its associated template in a standard PCR reaction. The oligonucleotide probes were employed in standard quantitative PCR amplification reactions with cDNA libraries isolated from different human tumor and normal human tissue samples and analyzed by agarose gel electrophoresis so as to obtain a quantitative determination of the level of expression of the PRO polypeptide-encoding nucleic acid in the various tumor and normal tissues tested. β -actin was used as a control to assure that equivalent amounts of nucleic acid was used in each reaction. Identification of the differential expression of the PRO polypeptide-encoding nucleic acid in one or more tumor tissues as compared to one or more normal tissues of the same tissue type renders the molecule useful diagnostically for the determination of the presence or absence of tumor in a subject suspected of possessing a tumor as well as therapeutically as a target for the treatment of a tumor in a subject possessing such a tumor. These assays provided the following results.

	<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
20	DNA26843-1389	normal lung rectum tumor	lung tumor normal rectum
25	DNA30867-1335	normal kidney	kidney tumor
30	DNA40621-1440	normal lung	lung tumor
35	DNA40625-1189	normal lung	lung tumor
40	DNA45409-2511	melanoma tumor	normal skin
45	DNA56406-1704	kidney tumor normal skin	normal kidney melanoma tumor
	DNA56410-1414	normal stomach	stomach tumor
	DNA56436-1448	normal skin	melanoma tumor
	DNA56855-1447	normal esophagus rectum tumor	esophageal tumor normal rectum
	DNA56860-1510	normal kidney rectum tumor	kidney tumor normal rectum
	DNA56862-1343	kidney tumor normal lung	normal kidney lung tumor

	<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
	DNA61755-1554	normal stomach kidney tumor	stomach tumor normal kidney
5	DNA62812-1594	normal stomach normal lung normal rectum normal skin	stomach tumor lung tumor rectum tumor melanoma tumor
10	DNA62815-1576	esophageal tumor	normal esophagus
	DNA64881-1602	normal stomach normal lung	stomach tumor lung tumor
15	DNA64902-1667	esophageal tumor kidney tumor	normal esophagus normal kidney
	DNA65403-1565	normal esophagus	esophageal tumor
20	DNA66308-1537	normal lung	lung tumor
	DNA66519-1535	kidney tumor	normal kidney
25	DNA66521-1583	normal esophagus normal stomach normal lung normal rectum normal skin	esophageal tumor stomach tumor lung tumor rectum tumor melanoma tumor
30	DNA66658-1584	normal lung melanoma tumor	lung tumor normal skin
	DNA66660-1585	lung tumor	normal lung
35	DNA66674-1599	kidney tumor normal lung	normal kidney lung tumor
	DNA68862-2546	melanoma tumor	normal skin
40	DNA68866-1644	normal stomach	stomach tumor
	DNA68871-1638	lung tumor normal skin	normal lung melanoma tumor
45	DNA68880-1676	normal lung normal skin	lung tumor melanoma tumor
	DNA68883-1691	esophageal tumor	normal esophagus
50	DNA68885-1678	lung tumor	normal lung
	DNA71277-1636	normal stomach	stomach tumor
	DNA73734-1680	normal lung	lung tumor

The assay is performed as follows. A PRO polypeptide of the present invention suspected of being a ligand for a receptor is expressed as a fusion protein containing the Fc domain of human IgG (an immunoadhesin). Receptor-ligand binding is detected by allowing interaction of the immunoadhesin polypeptide with cells (e.g. Cos cells) expressing candidate PRO polypeptide receptors and visualization of bound immunoadhesin with fluorescent reagents directed toward the Fc fusion domain and examination by microscope.

- 5 Cells expressing candidate receptors are produced by transient transfection, in parallel, of defined subsets of a library of cDNA expression vectors encoding PRO polypeptides that may function as receptor molecules. Cells are then incubated for 1 hour in the presence of the PRO polypeptide immunoadhesin being tested for possible receptor binding. The cells are then washed and fixed with paraformaldehyde. The cells are then incubated with fluorescent conjugated antibody directed against the Fc portion of the PRO polypeptide immunoadhesin (e.g. 10 FITC conjugated goat anti-human-Fc antibody). The cells are then washed again and examined by microscope. A positive interaction is judged by the presence of fluorescent labeling of cells transfected with cDNA encoding a particular PRO polypeptide receptor or pool of receptors and an absence of similar fluorescent labeling of similarly prepared cells that have been transfected with other cDNA or pools of cDNA. If a defined pool of 15 cDNA expression vectors is judged to be positive for interaction with a PRO polypeptide immunoadhesin, the individual cDNA species that comprise the pool are tested individually (the pool is "broken down") to determine the specific cDNA that encodes a receptor able to interact with the PRO polypeptide immunoadhesin.

In another embodiment of this assay, an epitope-tagged potential ligand PRO polypeptide (e.g. 8 histidine "His" tag) is allowed to interact with a panel of potential receptor PRO polypeptide molecules that have 20 been expressed as fusions with the Fc domain of human IgG (immunoadhesins). Following a 1 hour co-incubation with the epitope tagged PRO polypeptide, the candidate receptors are each immunoprecipitated with protein A beads and the beads are washed. Potential ligand interaction is determined by western blot analysis of the immunoprecipitated complexes with antibody directed towards the epitope tag. An interaction is judged to occur if a band of the anticipated molecular weight of the epitope tagged protein is observed in the 25 western blot analysis with a candidate receptor, but is not observed to occur with the other members of the panel of potential receptors.

Using these assays, the following receptor/ligand interactions have been herein identified:

- (1) PRO10272 binds to PRO5801.
20 (2) PRO20110 binds to the human IL-17 receptor (Yao et al., *Cytokine* 9(11):794-800 (1997); also herein designated as PRO1) and to PRO20040.
30 (3) PRO10096 binds to PRO20233.
(4) PRO19670 binds to PRO1890.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs 35 that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16),
5 Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54),
10 Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92),
15 Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 20 25 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166) and Figure 168 (SEQ ID NO:168).

2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure
30 35

Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165) and Figure 167 (SEQ ID NO:167).

4. Isolated nucleic acid having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

5

5. A vector comprising the nucleic acid of Claim 1.

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

10

7. A host cell comprising the vector of Claim 5.

8. The host cell of Claim 7, wherein said cell is a CHO cell.

15

9. The host cell of Claim 7, wherein said cell is an *E. coli*.

10. The host cell of Claim 7, wherein said cell is a yeast cell.

11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

12. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102

58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64),
Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID
NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ
ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88
(SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure
5 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102),
Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ
ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116),
Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ
ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130),
10 Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ
ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144),
Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ
ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158),
Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ
15 ID NO:166) or Figure 168 (SEQ ID NO:168), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2
(SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ
ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18
(SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure
20 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32),
Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID
NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ
ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56
(SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure
25 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70),
Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID
NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ
ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94
(SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100),
30 Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ
ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114),
Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ
ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128),
Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ
35 ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142),
Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ
ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156),

(SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102),
5 Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130),
10 Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ
15 ID NO:166) or Figure 168 (SEQ ID NO:168), lacking its associated signal peptide;

(b) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure

23. The method according to Claim 21, wherein said E, F, G, H or I polypeptide is labeled with a detectable label.

24. The method according to Claim 21, wherein said E, F, G, H or I polypeptide is attached to a solid support.

5

25. A method of detecting a polypeptide designated as E, F, G, H or I in a sample suspected of containing an E, F, G, H or I polypeptide, said method comprising contacting said sample with a polypeptide designated herein as A, B, C or D and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of an A, B, C or D polypeptide in said sample and wherein A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a PRO10096 polypeptide, D is a PRO19670 polypeptide, E is a PRO5801 polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a PRO20233 polypeptide and I is a PRO1890 polypeptide.

10

26. The method according to Claim 25, wherein said sample comprises cells suspected of expressing said E, F, G, H or I polypeptide.

15

27. The method according to Claim 25, wherein said A, B, C or D polypeptide is labeled with a detectable label.

20

28. The method according to Claim 25, wherein said A, B, C or D polypeptide is attached to a solid support.

25

29. A method of linking a bioactive molecule to a cell expressing a polypeptide designated as A, B, C or D, said method comprising contacting said cell with a polypeptide designated as E, F, G, H or I that is bound to said bioactive molecule and allowing said A, B, C or D and said E, F, G, H or I polypeptides to bind to one another, thereby linking said bioactive molecules to said cell, wherein A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a PRO10096 polypeptide, D is a PRO19670 polypeptide, E is a PRO5801 polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a PRO20233 polypeptide and I is a PRO1890 polypeptide.

30

30. The method according to Claim 29, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

35

31. The method according to Claim 29, wherein said bioactive molecule causes the death of said cell.

1/168

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGCTGGTAAGGATTACAAAAGGTGCAGGTA
TGAGCAGGTCTGAAGACTAACATTGTGAAGTTGAAACAGAAAACCTGTTAGAA**ATG**TGG
TGGTTTCAGCAAGGCCTCAGTTCCCTCAGCCCTGTAATTGGACATCTGCTGCTTTC
ATATTTCATACATTACTGCAGTAACACTCCACCATAAGACCCGGCTTACCTATATCAGT
GACACTGGTACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTT
TTATGCATTGCTACCATTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCAGTGAAGAGAAC
GTTATCATCAAATTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTCT
ATTGTGCAAACCTCCAGAAAACAACCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACC
TTTGGTATGGCCTATTATATGTTGTTCAGACCATCCTTACCAAATGCAGCCCCAAA
ATCCATGGCAAACAAGTCTCTGGATCAGACTGTTGTTATCTGGTGTGGAGTAAGTGCA
CTTAGCATGCTGACTTGCTCATCAGTTGCACAGTGGCAATTGGGACTGATTAGAACAG
AAACTCCATTGGAACCCCGAGGACAAAGCTTATCTGCTTCACATGATCACTACTGCAGCAGAA
TGGTCTATGTCATTTCCTTCTTGGTTTTCTGACTTACATTGTTGATTTCAGAAAATT
TCTTACGGGTGGAAGCCAATTACATGGATTAACCTCTATGACACTGCACCTGCCCTATT
AACAAATGAACGAACACGGCTACTTCCAGAGATATT**TGA**TGAAAGGATAAAATTTCTGAA
TGATTATGATTCTCAGGGATTGGGAAAGGTTCACAGAAGTTGCTTATTCTCTGAAATT
TCAACCACCTAATCAAGGCTGACAGTAACACTGATGAATGCTGATAATCAGGAAACATGAAAG
AAGCCATTGATAGATTATTCTAAAGGATATCATCAAGAAGACTATTAAAACACCTATGCCT
ATACTTTTATCTCAGAAAATAAGTCAAAAGACTATG

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGGAGAAGGC
AGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCCTGGGGAGTGGAGTGGAGGC
AGGAGCCTCCTACACTTCGCC**TG**GAGTTCTCATCGACTCCAGCATCATGATTACCTCCC
AGATACTATTTTGATTGGCTTCTCATGCGCCAATTGTTAAAGACTATGAGA
TACGTCAGTATGTTACAGGTGATCTCTCCGTGACGTTGCATTTCAGGACCATGTTG
AGCTCATCATCTTGAAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACTGGAAAAA
TGAACCTGTGTGAATTCTGCTGATCCTGGTTTACATGGTGCCTTTACATTGGCTATTTA
TTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTTCCGTCTTATGGCTGA
CCTTATGTATTCTTGAAACTAGGAGATCCCTTCCATTCTCAGCCAAAACATGGGA
TCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAGTGACTCTCATGGCTCTTC
TTTCTGGATTGGTGTCAACTGCCATACACTACATGTCTTACTCCTCAGGAATGTGA
CTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCAGGATATGATCATAAGCA
AAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTACCAACTCAGCATCAGGAAGTGAAAC
CATCAGGTTCTGGGAATGATAAAAAGTGTACCAACTCAGCATCAGGAAGTGAAAC
CTCTTATTCAACAGGAAGTGGATGCTTGGAGAATTAAGCAGGCAGCTTCTGGAAACAG
CTGATCTATATGCTACCAAGGAGAGAACAGAACACCTCAAGGGAAATATTTA
ATTTCTGGTACTTTCTATTTACTGTGTTGGAAAATTTCATGGTACCATCAATA
TTGTTTGATCGAGTTGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATT
ATCTGGGAATCCAATTGATGTGAAGTTGGTCCAACACATTCCATTCTTGTGGAA
TAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTATGCCATCT
CTAGCAGTAAGCCTCCAATGTCATTGCTGCTATTAGCACAGATAATGGCATGACTTTG
TCTCCTCTGTGCTGATCCGAATGAGTATGCCTTAGAATACCGCACCATAACTGAAG
TCCTGGAGAACACTGCAACTCTATCACCCTGGTGTGATCTCCTGGTCAGCG
CTCTCTAGCATACTCTCCTCTATTGGTCACAAACAGGCACCAGAGAACGAAATGGCAC
CT**TG**AACCTAACGCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTTAGATATAAGAGG
GGGGAAAATGGAACCAGGGCTGACATTATAACAAACAAAATGCTATGGTAGCATT
CACCTCATAGCATACTCTTCCCCGTCAAGGTGATACTATGACCATGAGTAGCATGCCAGA
ACATGAGAGGGAGAACACTAACAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGG
CTGGTGTAGAGGCCAGAGGCCAGAACACTAAAGGTGAAAATACACTGGAACCTGGG
AAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATITCGTTAAGGTTACATGGAAA
AGTTATAGCTTGCTTGAGATTGACTCATTAACAGAGACTGTAACAAAAAA
AAAAAAAAAGGCCAGCTGACTCTAGAGTCGACCTGCAGAACGTTGGCCGCATGGCCAACT
TGTATTGAGCTTATAATG

5/168

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACACATAGTT
CCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAAGTGGCCATCTG
AGGTGTTCCCTGGCTCTGAAGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGGTGCCT
CTCACTCCATCTGGACCACGAGGCTCCTGGCCAAGGCTTTGCGTGAGAAGAGCTTCC
ATCCAGGTGTATGCAGAATTATGGGGATCACCCCTGTGAGCAAAAAGGCAGACCAGCAGCTG
AATTTCACAGAACGTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTGGCCGGCAAGGACCAA
GTTGAAACAGCCTGAAAGCTAGCTTGAAACTTGCAGCTATGGCTGGGTTGGAGATGGATT
GTGGTCATCTCTAGGATTAGCCAAACCCAAAGTGTGGAAAAATGGGTGGGTGTCTGATT
TGGAAAGGTTCCAGTGAGCCGACAGTTGCAGCCTATTGTTACAACATCTGATACTGGACT
AACTCGTGCATTCCAGAAATTATCACCAACCAAGATCCCATTCAACACTCAAACGTGCAACA
CAAACAAACAGAACATTATTGTCAAGTACAGTACCTACTCGGTGGCATCCCCTACTCTACAATA
CCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAAATTG
ATTGTGTACAGAACGTTTATGGAAACTAGCACCATGTCTACAGAAACTGAACCATTGTT
GAAAATAAGCAGCATTCAAGAACGAGCTGCTGGGTTGGAGGTGCCCCACGGCTCTGCTA
GTGCTTGCTCTCCTCTTGGTGTGCTCAGCTGGTCTGGATTTGCTATGTCAAAAGGTAT
GTGAAGGCCTTCCCTTTACAAACAAGAACGAGAACGAGGAAATGATCGAAACCAAAGTAGTA
AAGGAGGAGAAGGCCAATGATAGAACCCATTAGAGGAATCAAAGAAAATGATAAAAACCCA
GAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCAGTCCTGGAGCTGAAGTT**TAGATGAGA**
CAGAAATGAGGAGACACACCTGAGGCTGGTTCTTCATGCTCCTTACCCCTGCCAGCTGGG
GAAATCAAAGGGCAAAGAACCAAAGAACGAGAACGAGGAAATGCCCTCCTTATTGTAACCCT
TCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAACGAGAACGAGGAAATGGAATCAGC
GTCTGGATCCTATCCTCTACCTCCAAAGCTCCCACGGCCTTCTAGCCTGGCTATGTCTA
ATAATATCCCACGGAGAACGGAGTTTGCAAAGTGCAAGGACCTAAACATCTCATCAGTA
TCCAGTGGTAAAAGGCCCTGGCTGTGAGGCTAGGGTAGGTTGAAAGCCAAGGAGTCAC
GAGACCAAGGCTTCTACTGATTCCGAGCTCAGACCCCTTCTCAGCTCTGAAAGAGAAA
CACGTATCCCACCTGACATGTCCCTCTGAGCCGGTAAGAGAACGAGGAAATGGCAGAAAAGTT
AGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAAGCTAAAATAAA
GAAATAGAACAAAGCTGAGGATACGACAGTACACTGTAGCAGGGACTGTAACACAGACAGG
GTCAAAGTGTGTTCTGAACACATTGAGGTTGAATCACTGTTAGAACACACACTTACTT
TTTCTGGTCTCTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACAAAAT
AAAAACTCTATAAATTCTATTGAGTTACAGAAATGATTACTAAGGAAGATTACT
CAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTCA
AGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCTCAAAAATTGCACATAGTA
GAACGCTATCTGGGAAGCTATTGTTCTGAGTTGATATTCTAGCTTATCTACTTCAAAC
AATTTTATTGCTGAGACTAATCTTATTCTAATATGGCAACCATTATAACCT
TAATTATTATAACATACCTAACAGAAGTACATTGTTACCTCTATATACCAAAAGCACATTAA
AAGTGCCATTAACAAATGTACTAGCCCTCCTTTCCAACAAAGAAGGGACTGAGAGATGC
AGAAATATTGACAAAAATTAAAGCATTAGAAAACCTT

7/168

FIGURE 7

CGCCGCGCTCCCGCACCGCGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCCG
GC GG CCT CCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGG
GC GG CGG CT GC GGG CG CAG AG CG GAG ATG CAG CG GCT TGGG CAC CCT GCT GT GC CT GCT GC
TGGCGGGCGGGTCCCCACGGCCCCCGCCCGCTCCGACGGCAGCCTCGCTCCAGTCAAGC
CCGGCCCCGGCTCTCAGCTACCCGAGGAGGAGGACCCCTCAATGAGATGTCCGCAGGTTG
AGGAAGTGTGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAG
AAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACATTACCTCCCAGCTATACAATG
AGACCAACACAGACACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAATTACAAGA
TAACCAACAACCAGACTGGACAAATGGTCTTCAGAGACAGTTATCACATCTGTGGGAGACG
AAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCC
AGTTGCCAGCTCCAGTACACCTGCCAGCCATGCCGGGGCAGAGGATGCTCTGCACCCGGG
ACAGTAGTGTGGAGACCAGCTGTGTCTGGGTCAGTGCACACCCCTGCCGTGGAGGGCAGCTTGCAC
GCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGCTGTGCTGTGCCTTCC
AGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCGTGGAGGGCAGCTTGCAC
CCGCCAGCCGGCTCTGGACCTCATCACCTGGAGCTAGAGCCTGATGGAGCCTGGACCGAT
GCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTGGTGTATGTGTGCAAGC
CGACCTCGTGGGAGCCGTGACCAAGATGGGAGATCCTGCTGCCAGAGAGGTCCCCGATG
AGTATGAAGTTGGCAGCTCATGGAGGAGGTGCCAGGGCTGGAGGACCTGGAGAGGAGC
TGACTGAAGAGATGGCGCTGGGGAGCCTGCCGTGCCCTGCACACTGCTGGAGGGGAAG
AGATT TAG AT CT GG ACC AGG CT GT GG G TAG AT GT G CA A T AG A A AT AG C T A A T T A T T C C C C A
GGTGTGTGCTTAGGCGTGGCTGACCAAGGCTTCTCCTACATCTCTCCAGTAAGTTCC
CCTCTGGCTTGACAGCATGAGGTGTTGTGCATTGTCAGCTCCCCAGGCTGTTCTCCAGGC
TTCACAGTCTGGCCTGGAGAGTCAGGCAGGGTTAAACTGCAGGAGCAGTTGCCACCCCT
GTCCAGATTATTGGCTGCTTGCCTCTACAGTTGGCAGACAGCCGTTGTTCTACATGGCTT
TGATAATTGTTGAGGGAGGAGATGGAAACAATGTGGAGTCTCCTCTGATTGGTTTGGGG
AAATGTGGAGAAGAGTGCCTGTTGCAAACATCAACCTGGCAAAAATGCAACAAATGAATT
TTCCACGCAGTTCTTCCATGGCATAGGTAAAGCTGTGCCTCAGCTGTTGCAGATGAAATGT
TCTGTTCACCCCTGCATTACATGTGTTATTACATCCAGCAGTGTGCTCAGCTCCTACCTCTGT
GCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGGAGGGGG
CATTGTTCTCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCCAAGT
CACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTTGTGACTCTAACGCTCAGTGCCTCTC
CACTACCCACACCAGCCTGGTGCCACCAAAAGTGTCTCCAAAAGGAAGGAGAATGGGATT
TTTCTTGAGGCATGCACATCTGGAATTAAAGGTCAAACATAATTCTCACATCCCTCTAAAGTA
ACTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGCTTCTAATGAAGACAAT
GATATTGACACTGTCCCTCTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCG
TAGCATACAGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAATGA
AATTGCAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACC
GAGCAGGGCTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTACGCCACTC
CACAAATGATGTTTCAGGTGTCATGGACTGTTGCCACCATGTATTCTACAGAGTCTTAAA
GTTAAAGTTGCACATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACAT
AAGTTGCATTAGAAATCAAGCATAAATCACTCAACTGCAAAAAAAAAAAAAAAA
AAA

9/168

FIGURE 9

CGGACGC GTGGCGGACCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCC
ACGGCCCACCTTGTGAACTCCTCGTGC CCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCC
AAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGCCTGGGGCTCTTCTGG
ACCCTTAACTGGGTACTGCCCTGGCCAATGCGCCTCGCTGGAGCCTTGCCTCCTCTAC
TGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTGCCTTCATCCGCACA
CTCCGTTACCACACTGGTCATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCC
CGGGTCATCTGGAGTATTGACCAAGCTCAGAGGAGTGCAGAACCTGTAGCCCGCTGC
ATCATGTGCTGTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTTCCTAAACCGC
AATGCATACATCATGATGCCATCTACGGGAAGAATTCTGTGTCTCAGCCAAAATGCGTTC
ATGCTACTCATGCGAACATTGTCAGGGTGGTCGTGGAGGCCTGGGGCTCTGTCTTTCTCCGGT
CGCATCCCGGGCTGGTAAAGACTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATC
ATGACCTCCATCCTGGGGCTATGTCATGCCAGCGGCTTCTCAGCGTTTCGGCATGTGT
GTGGACACGCTTCTGCTTGGAAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGG
CCCTACTACATGTCCAAGAGCCTCTAAAGATTCTGGCAAGAAGAACGAGGCGCCCCGGAC
AACAAAGAGAGGAAGTGACAGCTCCGGCCCTGATCCAGGACTGCACCCACCCCCACCGT
CCAGCCATCCAACCTCACTCGCCTTACAGGTCTCCATTGTGGTAAAAAAAGTTTAGGC
CAGGCGCCGTGGCTACGCCTGTAATCCAACACTTGGAGAGGCTGAGGCAGGGATCACCTG
AGTCAGGAGTTGAGACCAGCCTGGCAACATGGTAAACCTCCGTCTCTATTAAAAATCAA
AAATTAGCCGAGAGTGGTGGCATGCACCTGTCTCCAGCTACTCGGGAGGCTGAGGCAGGAG
AATCGCTTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAACC
TGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAAAGATTATTAAAGATATT
TGTAACTC

11/168

FIGURE 11

GCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCC**ATG**GGGGCCTGCCTGGAG
CCTGCTCCCTGCTCAGCTGCCGTCTGCCTCTGCCGCTCTGCCCTGCATCCTGTGCAGCT
GCTGCCCGCCAGCCGCAACTCCACCCTGAGGCCCTCATCTCACGTTCTCCTCTCCTGG
GGGTGCTGGTGTCCATCATTATGCTGAGCCCAGGGCTGGAGAGTCAGCTCTACAAGCTGCCCT
GGGTGTGTGAGGAGGGGCCGGATCCCCACCGTCCTGCAGGCCACATCGACTGTGGCTCCC
TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTCGCCACGGCGGCCCTCTTCTTCTTCTT
TCACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGACCCCCGGGCTGCCATCCAGAATGGGT
TTTGGTTCTTAAGTTCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACG
GCTCCTCACCAACATCTGGTTCTACTTCGGCGTGTGGCTCCTCCTCTTCTACATCCCTCATCC
AGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGCAAGGCCGAGG
AGTGCGATTCCCGTGCCTGGTACGCAGGCCTTTCTTCTTCACTCCTCTTCACTTGCTGT
CGATCGGGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGCGGCTGCCACGAGGGCA
AGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGTCCTGCCA
AGGTCCAGGACGCCAACCTGGGCTGCTGCAGGCCTCGGTACCCCTCTACACCA
TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAA
CCCAGCTGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATG
CCCCGAGCATTGTGGCCTCATCATCTTCTCCTGTGCACCCCTTTCATCAGTCTGCGCTCCT
CAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCA
CACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGCCTTGACAACGAGCAGGACG
GCGTCACCTACAGCTACTCCTTCCACTTCTGCCTGGTGTGGCTCCTCACTGCACGTCTGA
TGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGATGATCAGCACGTGGACCGCCG
TGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCTCCTCTACCTGTGGACCCCTGGTAGCCC
CACTCCTCCTGCGCAACCGCGACTTCAGC**TGA**GGCAGCCTCACAGCCTGCCATCTGGTGCCTC
CTGCCACCTGGTGCCTCTGGCTCGGTGACAGCCAACCTGCCCTCCCCACACCAATCAGCC
AGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGGCTTCTAGTCGT
AGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCACACCCAC
ACGGTGGAGCTGCCTCTTCTTCCCTCCCTGGTGTGGCCATACTCAGCATCTGGATGAAA
GGGCTCCCTTGTCTCAGGCTCACGGAGCGGGCTGCTGGAGAGAGAGCGGGGAACTCCCACC
ACAGTGGGCACTCCGGCACTGAAGCCCTGGTGTTCCTGGTACGTCCCCCAGGGACCCCTGCC
CCCTTCTGGACTTCGTGCTTACTGAGTCTCTAAGACTTTCTAATAAAACAAGCCAGTGC
TGTAAAAAAA

13/168

FIGURE 13

CGGGCCAGCCTGGGCGGCCAGGAACCACCGTTAAGGTGTCTCTCTTAGGGATGGT
GAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGG**ATG**AACCACCTGCCAGAAGACATGGAG
AACGCTCTCACCGGGAGCCAGAGCTCCCATGCTCTCTGCAGCAATATCCATTCCATCAACCCCC
ACACAACATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC
AGGAGGACTTCTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATA
GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC
TATTCTCATATTTGATATATTCTCTGGCAGTTTCGATTAAAGTGTAAATACTGCA
TATGCTGTGTGCAGACTCGGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC
TTTTTACTAGCAAAAGTGATCCTTCGAAGCTTTCTCAAGGGCTTTGGCTATGTGCTG
CCCATCATTTCATTTCATCCTTGCCCTGGATTGAGACGTGGTCCTGGATTCAAAGTGTACCT
CAAGAACAGAAGAACAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTT
ATACCTGGTGGCTTCTGATGGCAGTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAA
GAAGCTGAAGAAAACAGGACAGTGAGAACCACTTTAGAAACT**TGA**GTACTACTTTGTTA
AATGTGAAAAACCCCTCACAGAAAGTCATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTG
TCGACAGTAAAGTTGAAATGGTGACGTCCACTGCTGGCTTATTGAACAGCTAATAAGATT
ATTTATTGTAATACCTCACAAACGTTGTAACCATATCCATGCACATTAGTTGCCTGCCTGTGG
CTGGTAAGGTAATGTCATGATTCATCCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAAA
TAGGTGAAGAAAGTCTTGTGCTGTATTCTTAATCAAAAGACTTAATATATTGAAGTAACACTT
TTTAGTAAGCAAGACACCTTTATTCAATTACAGAACGAAATTGGAATTGTTTGTCTGTT
CAGATTATTGTTGATTCTTCTTTAACACTCTACATTCCCTGTTTTAACATGCACAA
TGTGCTTTGTACAGTTAAAAAGTGTAAATAAAATCTGACATGTCAATGTGGCTAGTTA
TTTTCTTGTGTTGCATTATGTGATGGCCTGAAGTGTGGACTTGCAAAAGGGAAAGAAAGG
AATTGCGAACATGTAAAATGTCACCAGACATTGTATTATTTATCATGAAATCATGTT
TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGCACAAATGACTAAACC
ATTCATATCATGTTCTTGCCTTCAGCCAATTCAATTAAATGAACTAAATTAAAAA

15/168

FIGURE 15

ACTCGAACGCAGTTGCTCGGGACCCAGGACCCCCCTCGGGCCCAGCCGCCAGGAAAGACTGA
GGCCGCGGCCTGCCCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAG**ATGTGCT**
CCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGTGCAGGGCT
GCCCATCCGGCTGCCAGTGCAGCCACAGACAGTCTTCTGCACTGCCGCCAGGGGACCA
CGGTGCCCGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATCACCA
TGCTCGACGCAGGCAGCTTGCCGGCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAAC
AGATCGCCAGCCTGCCAGCGGGCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGA
CGGCCAACAGGCTGCATGAAATACCAATGAGACCTTCCGTGGCCTGCGGCCCTGAGCGCC
TCTACCTGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACC
TCCGGAGCTCAAGCTGCAGGACAACAGAGCTGCAGGGCACTGCCCGCTGCCCTGCCCG
TGCTGCTGCTGGACCTCAGCCACAACAGCCTCTGGCCCTGGAGCCGGCATCCTGGACACTG
CCAACGTGGAGGCGCTGCCGTGGCTGGCTGGGCTGCAGCAGCTGGACGAGGGGCTTTCA
GCCGCTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAGCGAGTGC
ACCTTGATCCGAGGCCTCCGGGGCCTGACGCCCTGCCGGCAACACCCGATTGCCAGC
TGCAGGCCGAGGACCTGCCGGCCTGGCTGCCCTGCAGGAGCTGGATGTGAGAACCTAAGCC
TGCAGGCCCTGCCGTGGCACCTCTGCCCTCTTCCCCGCCCTGCCGTGCTGGCAGCTGCC
GCAACCCCTTCAACTGCCGTGCCCCCTGAGCTGGTTGGCCCTGGGTGCCGAGAGCCACG
TCACACTGCCAGCCCTGAGGAGACGCCGTGCCACTTCCGCCAAGAACGCTGCCGGCTGC
TCCGGAGCTTGACTACGCCGACTTGGCTGCCAGCCACCACACCAGCCACAGTGCCCA
CCACGAGGCCGTGGTGCCGGAGCCCACAGCCTGTCTTAGCTTGGCTCTACCTGGCTTA
GCCAACAGGCCGGCACTGAGGCCCAAGCCCTCCACTGCCCAACCGACTGTAGGGC
CTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCCTAACATGGGGCACATGCCAC
GGACACGGCACCACTGGCGTGGCTGTGCCCTGAAGGCTTACGGGCTGTACTGTGAGAGCC
AGATGGGGCAGGGACACGGCCAGCCCTACACCAAGTCACGCCGAGGCCACACGGCT
CCCTGGGCATGGAGCCGGTGAGGCCACCTCCCTGCCGTGGCTGCCGCTACCTCCAGG
GGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATGCCAACCTATGCCCTGATAAGC
GGCTGGTGACGCTGCGACTGCCCTGCCCTCGCTGAGTACACGGTACCCAGCTGCC
ACGCCACTTACCCGTGTGTCTAGCCTTGGGCCGGGCGGGTGCCGGAGGGCGAGGAGG
CCTGCCGGGAGGCCATACACCCCCAGCCGTCCACTCCAACCACGCCCAAGTCACCCAGGCC
GCGAGGGCAACTGCCGTCTCATTGCCCGCCCTGCCGCCGGTGCCTCTGCCCGCTGG
CTGCCGTGGGGCAGCTACTGTGTGCCGGGGGGCATGGCAGCAGCGCTCAGGACA
AAGGGCAGGTGGGCCAGGGCTGGCCCTGGAACTGGAGGGAGTGAAGGTCCCTGGAGC
CAGGCCGAAGGCAACAGAGGCCGTGGAGAGGCCCTGCCAGCAGGTCTGAGTGTGAGGTG
CACTCATGGCCTCCAGGCCCTGCCCTCCAGTCACCCCTCCACGCAAAGCCCTACAT**TAA**
CCAGAGAGAGACAGGGCAGCTGGGCCGGCTCTCAGCCAGTGAGATGCCAGGCC
CTGCCACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCT
CACAGCTGGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGT
TGACGAGCCCTAACGTCCCCAGAACCGAGTGCTATGAGGACAGTGTCCGCC
AACGTGCAGTCCCTGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCC
CTGGCTCTCCCACTGCCAGGCGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAG
GGAGAGCGGGTAGGCGGCTGTGACTCTAGTCTTGGCCCCAGGAAGCAGAACAAAAGAA
ACTGGAAAGGAAGATGCTTGTAGAACATGTTTGCTTTAAAATATATATTTATAAGAG
ATCCTTCCCATTATTCTGGAAAGATGTTTCAAACTCAGAGACAAGGACTTTGGTT
TAAGACAAACGATGATGAAGGCCTTGTAGAAAAAATAAAAGATGAAGTGTGAAA

17/168

FIGURE 17

GCAGCGGCGAGGC GGCGGTGGCTGAGTC CGTGGCAGAGCGAAGGCACAGCTC**ATG**
CGGGTCCGGATAAGGCTGACGCTGCTGTGCGGTGCTGAGCTTGCCTCGCGTCC
TCGGATGAAGAACGGCAGGCCAGGATGAATCCTAGATTCAAGACTACTTGACATCAGATGAG
TCAGTAAAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTGAGAA
GAATCTGAATTAGAATCCTCTATTCAAGAACAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAA
AGTGTACAGAACAGATATCAGCTTCTAGAGTCTCCAATCCAGAAAACAAGGACTATGAAGAG
CCAAAGAAAGTACGGAAACCAGCTTGACGCCATTGAAGGCACAGCACATGGGGAGCCCTGC
CACTTCCCTTTCTTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGAT
GGCAGACTGTGGTGTGCTACAA CCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTGTGAA
ACTGAAGAACAGAGGCTGCTAACAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAAC TGGAA
ATGAAAATCCTTAATGGAAGCAATAAGAAAAGCCAAAAAGAGAACATATCGGTATCTCAA
AAGGCAGCAAGCATGAACCATAACCAAAAGCCCTGGAGAGAGTGTATGCTCTTTATTGGT
GATTACTTGCCACAGAACATCCAGGCAGCGAGAGAGATGTTGAGAACGCTGACTGAGGAAGGC
TCTCCCAAGGGACAGACTGCTCTGGCTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGT
CAGGCAAAGGCTTTGTATATTACATTGGAGCTCTGGGGCAATCTAACAGCCCACATG
GTTTGTTGTAAGTAGACTT**TAGT**GGAAAGGCTAATAATATTAAACATCAGAACAGAATTGTTGTTA
TAGCGGCCACAACCTTTCAGCTTCTGATCCAGATTGCTGTATTAAGACCAAAATTCA
GTTGAACCTCCTCAAATTCTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAG
TTGGTGGAGTCCACAATTCTTCTTAAATGATTAGTTGGCTGATTGCCCTAAAAGAGAG
ATCTGATAAAATGGCTCTTTAAATTCTCTGAGTTGGATTGT CAGAACATCATTACAT
TAGATTATCATAATTAAAAATTCTTAGTTCAAAATTGTAAATGGTGGCTATA
GAAAAAACACATGAAATATTACAAATTTGCAACAAATGCCCTAACAGAATTGTTAAAATTCA
TGGAGTTATTGTGCAGAACATGACTCCAGAGAGCTACTTCTGTTTACTTTCTGATT
GGCTGTCTCCCATTCTGGCATTGCTAGTGACACTGTGCCTGCTCCAGTAGTC
TCATTTCCTATTGCTAATTGTTACTTTCTTGCTAATTGGAAGATTAACTCATT
TTAATAAAATTATGTCTAAGATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

19/168

FIGURE 19

AATT CAG ATT TAAG CCC ATT CTG CAG TGG ATT CAT GAA CT AG CA AG AG GAC ACC AT CCT TC
TT GT ATT ATA CA AG AA AG GAG TGT ACCT ATC ACAC ACAG GGGG AAAA **ATG** CT TT GG GT GC
TAGG CCT CCT AAT CCT CT GT GG TTT CT GT GG ACT CGT AA AGG AAAA CT AA AG ATT GA AG AC A
TC ACT GATA AGT AC ATT TT ATC ACT GG AT GT GACT CGG CTT GG AA ACT TG GC AG CC AG AA
CT TT GATA AAAA AGG ATT CAT GT AAT CG CT GC CT GT CT GACT GA AT CAG GAT CA AC AG CT
TAA AGG CAG AA AC CT CAG AG AG ACT CGT ACT GT GCT T CT GG AT GT GACC AG CC AG AG AAT G
TCA AG AGG ACT GCCC AGT GGGT GA AGA ACCA AGT TGGG AG AA AG GT CT TGGG GT CT GAT CA
ATA AT GCT GGT GT TCC CGG GT CT GG CT CC ACT GACT GG CT GAC ACT AG AGG ACT AC AG AG
AAC CT ATT GA AGT GA AC CT GT TT GG ACT CAT CAG TGT GAC ACT AA AT AT GCT T CTT GG TCA
AG AA AG CT CA AGG GAG AG TT ATT AAT GT CT CC ACT GT TGG AG GT CG CTT GCA AT CG TT GG AG
GGG CT AT ACT CC AT CCA AAT AT GCA GT GGA AGG TT CA AT GAC AG CT TA AG AC GGG AC AT GA
AAG CT TT GG GT GC AC GT CT CAT GC AT TA GCA ACC AGG ATT GT CAAA AC AA ACT TG GC AG AT C
CAG TAA AG GT AAT GAAAAAA ACT CG CC ATT TGG AG CAG CT GT CT CC AG AC AT CAA AC AA AC
AA TAT GG AGA AG GT TAC ATT GAAAAAA AGT CT AG AC AA ACT GAA AGG CA AT AA AT CCT AT GT GA
AC AT GG AC CT CT CC GG GT AG AGT GC AT GG ACC AC GCT CT AACA AGT CT CT CC CT AAGA
CT C ATT AT GCG CT GG AAAA AG AT GC CAAA ATT T CT GG AT AC CT GT CT CA AT GCG AG CAG
CT TT GCA AG AC TT TT ATT GT TGA AAC AGA AAC AG CAG AG GCT GG CT AAT CCC AAGG CAG TG **TGA**
TC AG CT AACC ACA AAT GT CT CC CT CC AGG CT AT GAA ATT GCC GATT CA AG AAC AC AT CT C CT
TTT CA ACC CC ATT C CT TAT CT GCT CC A AC CT GG ACT C ATT TAG AT CGT GCT T ATT GG ATT GC
AAA AGG GAG TCC ACC AT CG CT GG GT AT CCC AGG GT CC CT GCT CA AG TTT CT TGG AAA AG
GAG GG CT GG AAT GG TAC AT CAC AT AGG CA AGT CCT GCC CT GT ATT TAGG CT T GCG CT GT GG
TGT GAT GT AAGG GAA ATT GAA AG ACT T GCC C ATT CAA AAT GAT CT T ACC GT GG C CT G C C C A
TG CT T ATT GGT CCCC AG C ATT AC AGT AACT GT GA AT GT TAA GT AT C AT CT CT T ATT C TAA AATA
TT AAA AG AT AAGT CA ACC CAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA

21/168

FIGURE 21

CTGAGGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGG
CGCACTCGCTTCCAGCACCTAACACACGGACTCGGACACGGAAAGGTTTCTTCTGGGAAGT
AAAAGGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAAC
AATTGACATTCAAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCAG
AGTAAATGAGCAAGCACTGAAGAAAATATTCAAATGTCAAAAGAATGTGGTAGGTTGGTA
CAAATTCCGTCGTCAATTCAAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAACTT
GCAGGAGCATTTCAAACCAAGACCTGTTCTGCTATTACACCAAGTATAAACAGA
AAGCTGCTCTACTCATCGACTGGAACATTCTTATATAAACCTCAAAAGGACTTTCACAG
GGTACCTTAGTGGTTGCCAATCTGGCATGCTGAACAACACTGGGTATAAAACTGTATCAGG
TTCCTGTATGTCCACTGGTTAGCCGAGCAGTACAAACACAGCTCTAAATTTTGAGAAGA
AGATGGATCCTAAAGGAGGTACATAAGATAATGAAATGTATGCTTCATTACAAGAGGAATT
AAAGAGTATATGCAAAAAGTGGAAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGT
AAACAGATTAAACGAGAAATTGAGAAAAGGGAGAGCAGATTCAAGGAGCAAGAGAGAA
GAACATCCAAAAGACCCCTCAGGAGAACATTTCATTGTCAGGCATTACGGACCTTTCC
AAATTCTGAATTCTTCATTCATGTGTTATGTCTTAAAAATAGACATGTTCTAAAGTAG
CTGTAACCTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAATGGTAGAACACACTGA
CATTCCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCATAAAGCCTAGACTAGA
TGACAGATGGCAATTCAAGAGATCTGGTTAGATAACACAAGACAAACGATCTAAAGCAA
TACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCCCAGAAACAGATGAAGAAAT
TGAAAAGATGAAGGGTTTGGTGAATATTCA CGGTCTCCTACATT**TGA**TCCTTTAACCTTA
CAAGGAGATTTTTATTGGCTGATGGTAAAGCAAACATTCTATTGTTACTATGTT
GAGCTACTTGCAGTAAGTCATTGTTACTATGTTCACCTGTTGAGTAATACACAGAT
AACTCTTAGTGCATTACTTCACAAAGTACTTTCAAACATCAGATGCTTTATTCCAAAC
CTTTTTTCACCTTCACTAAGTTGAGGGAAAGGCTTACACAGACACATTCTTAGAATT
GGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCAGCTAGGAAAGACAAGTC
AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATTGAGACCATGTCTA
TTAAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAAGAAATTATAT
GAAAATTATCTGAGTCATTAAATTCTCCTTAAGTGATACTTTTAGAAGTACATTATGGC
TAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTGCAAAACATCATCTAAATTT
AAAAAAAAAAAAAA

23/168

FIGURE 23

GGCACAGCCGCGCGCGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC
CAAGCAGCGCGCAGCGAACGCCCGCCGCCACACCCCTCTGCCGTCCCCGCCCTGCCACCCCTCC
TTCCCCCGCGTCCCCGCCTGCCCGCCAGTCAGCTTGCCGGGTCGCTGCCCGCGAAACCCCGAGGTACCCAGCC
CGCGCCTCTGCTTCCCTGGGCCGCGCCGCCTCCACGCCCTCCTCTCCCTGCCCTGGGCCGCGCTGGCACCGGGG
ACCGTTGCCCTGACCGAGGCCAGCTACTTTGCCCGCGTCTCCCGCCTGCCCTCTCCACCAACT
CCAACTCCTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTGGGCCGCTGCCGTAGCGCCGC
TTCCCCGTCCGGTCCCAAAGGTGGGAACCGTCCGCCCGGCCGACCATTGGCACGGTTGCCGTGCCGCCGCTT
CTCTGCACCTGGCAGTGCTCAGGCCCGCTGCTGGCTGCCGAGCTCAAGTCAGTCAGTCAGTCAGTCAGTCAGTC
CGTCTTACGTGTCCAAGGCTTCAACAAGAACGATGCCCTCCACAGAGATCAACGGTGTACATTGAAGATC
TCTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAA
AGTGTGGTCAGCGAACAGTCAATTTGCAAGCTGTCTTGCTTACGTTACAAGAACGTTGATGATTCTTC
AAAGAACACTTGTAAAATGCAGAGAAATCCCTGAATGATATGTTGTGAAGACATATGCCATTATACATGCAA
AATTCTGAGCTATTAAAGATCTCTCGTAGAGTTGAAACGTTACTACGTGGTGGAAATGTGAACCTGGAAAGAA
ATGCTAAATGACTCTGGCTCGCCTCTGGAGCGGATGTTCCGCCTGGTAACCTCCAGTACCAACTTACAGAT
GAGTATCTGGAATGTGTGAGCAAGTATACTGGAGCAGCTGAAGCCCTCGGAGATGTCCTCGCAAATTGAAGCTC
CAGGTTACTCGTCTTGTAGCAGCCGTACTTCGCTCAAGGTTAGCGGTTGCCGGAGATGTCGTGAGCAAG
GTCTCCGTGTAACCCCCACAGCCCAGTGTACCCATGCCCTGGTAAGAGATGATCTACTGCTCCACTGCCGGGGT
CTCGTGAUTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGCTGTTGCCAACCAAGGGGATCTC
GATTTGAATGGAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGCTTTCAACATTGAA
TCGGTCATGGATCCCATCGATGTGAAGATTCTGATGCTATTATGAACATGCAAGATAATAGTGTCAAGTGTCT
CAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCCTCCAGCTGGACGAATTCTCGTCCATCTGAAAGT
GCCCTCAGTGTGCTTCAAGACCACATCCCCGAGGAACGCCAACACAGCAGCTGGCACTAGTTGGACCGA
CTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAATTCTGGCCTCCCTCCGAGCAACGTTGCAAC
GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGAAAGGCAAAAGCAGGTACCTGTT
GCAGTGACAGGAATGGATTAGCCAACCAGGGCAACAACCAGAGGTTCCAGGTTGACACCAGCAAACCCAGACATA
CTGATCTCGTCAAATCATGGCTCTCGAGTGTGACCGAGCAAGATGAAGAACATGCAACATGGAAACGACGTG
GACTTCTTGTATCAGTGTGAAAGTACTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCCTCA
GAGTTGACTACAATGCCACTGACCATGCTGGAGAGTGCACATGAGAAAGCCGACAGTGTGGTGTCCGTCT
GGGGCACAGGCCTACCTCCTCACTGTCTCTGCATCTGGTCTGGTATGCAGAGAGAGTGGAGATAATTCTCA
AACTCTGAGAAAAAGTGTTCATCAAAAGTAAAGCACCAGTTATCACTTTCTACCATCCTAGTGAATTG
TTTTAAATGAATGGACAACAATGTAAGTTTACTATGTGGCACTGGTTAAGAAGTGTGACTTTGTTTCT
TCATTCACTTTGGAGGAAAAGGGACTGTGCATTGAGTTGGTCTGTGCTGCCAACCCATGTTAACATG
AACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTGTGATTATCACTTATTGTTGTATGTTT
TTCTCATTTCTGGTGTGGTTTTCTCAACTGTGATCTGCCCTGTTCTTACAAGCAAACCAGGGCT
CTTGGCACGTAACATGTACGTATTCTGAAATATTAAATAGCTGTACAGAACGAGGTTATTATCATGTTATC
TTATTAAAAAGAAAAAGCCCAAAAGC

25/168

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTG
ACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATTCCTCAAGC
AACTTACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGC
CACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGG
ACCGAGGCCAGGCTTAGGAGATGGCTCCAGGAAGGCAGCCAAGAATGTGAGTGCAAAGATT
GGTTCCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCC
CCTGTGATCATTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAA
ACAAGCATTCCAGAGCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGC
CTTGTAGGAGCTCTGAGCGCCACTCTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAG
CACACCTACCAGACACTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCAT
TCCAGTGCTCTCAAAAAGCATGTTTCAAGATCATTGTTGCTCTCTAGTGTCTT
CTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCTACCCAGGCTTAGGCTTAATTACCTGAAA
GATTCCAGGAAACTGTAGCTTCCTAGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTA
GCAAACAGAAGTCAATAATTTAAATGTCAAAAAAAAAAAAAAA

27/168

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGC
CCCCAGCGGAAGCACAGCTCAGAGCTGGTCTGCCATGGCT
CCCACCTGCACCATGGCTCTGGCTGGCAGCCCCGTGCA
AAAGCTACTTCCCTACCTGATGGAGAGCAAGAACGGAGCTCT
TCAGCCAGATAAAGGGCTTACAGGAGCCTCCGGAAAGTGG
CCCTACTGGAGCTGGAGCTGGAGCTGGAGCTGGGCTGCG
GAACCGGAGCCAACTTCAGTTCTACCCACCGGGCTGC
AGGGTCACCTGCCTAGACCCAAATC
CCCACTTGAGAAGTTCTGACAAAGAGCATGGCTGAGA
ACAGGCACCTCCAATATGAGCGGT
TTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGG
CTGATGGCTCCATGGATGTGGTGGTCT
GCACTCTGGTGCCTGTGCAGAGCCAAGGAAGGT
CCTGCAGGGAGGTCCGGAGAGTAC
TGAGACCGGGAGGTGTGCTCTTTCTGGAG
GCATGTGGCAGCAAGTTTCGAGCCCACCTGG
AAACACATTGGGATGGCTGCTGCCTCACCA
GAGAGACCTGGAAGGATCTTGAGAACGCC
AGTTCTCCGAAATCCAATGGAAGCAG
ACAGCCCCCTCC
CTCCCTGAAGTGGCTACCTGTTGGGCC
CACATCATGGAAAGGCTGTCAA
ACAATCTTCC
CAAGCTCCAAGGC
ACTCATTGCTC
CTTCCCCAAC
CTGC
AGGGCAATCT
TA
ACTTCA
ATCCC
GCCT
CGACAGTGA
AAAAGCT
CTACTTCT
ACGCT
GACCC
AGGGAG
AAACACT
TAGGACC
CTGTT
GTATC
CTCA
ACTG
CAAGTTCT
GGACT
AGTCT
CCCA
ACGTT
GCCT
CCA
ATGTT
GTCC
CTT
CGTT
CC
GGTAAAGCT
CCT
CTCG
CTT
CCT
GAGG
CTAC
ACCC
ATGCG
TCT
TAGGA
ACTGG
TCACAA
AAGTC
ATGGT
GCCT
GC
ATCC
CTGCC
AAAG
CCCC
CTGAC
CCCT
CT
CCCC
ACT
ACC
AC
CT
CT
CCT
GAG
CTGGGG
GATGCC
AGAG
CAAG
ACT
CAA
AGAG
GC
AGAG
GTT
CT
CAA
AT
TTTT
AATA
AGAC
GAA
ACC
ACG

29/168

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTG
CTGCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTGCCCAGAAAGTCTCT
TCTGCCACTGACGCCCATCAGGGATTGGGCCTCTTCCCCCTCCTGTGTCTCCTG
CCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGG
GATGGCTAAGAAAGCTGGAGATAGGAAACAGAAGAGGGTAGTGGGTGGCTAGGGGGCTGC
CTTATTAAAGTGGTTGTTATGATTCTTATACTAATTATAACAAAGATATTAAGGCCCTGTT
CATTAAGAAATTGTTCCCTCCCTGTGTTCAATGTTGTAAGATTGTTCTGTGTAAATATG
TCTTTATAATAAACAGTAAAAGCTGAAAAA

31/168

FIGURE 31

GTTCGAATTCTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAG
TTCCCTCCAAGCAAGTCATTCCCTTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACT
CCCTATTCGCATCTGTTGATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTCCAATCA
TGTCGGGAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGCCCCATGA
TGTTTACCTTCAGATTCATCACCAACCCCTCTGGTCACATTTCATTCATTGGTTATTTGG
GATTGTTGTTGTCTGCGGTGTTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCA
TAGAATTGGACACAGAAAGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAG
GCATCACGGCAGTGCCTCGTCTGATTTCAGAAAGAGAATAAAATTGACAGTTG
AGCTTTCCAATCACAAATAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGT
GGACATTGCCATCCTCATTTCTGGTCCTCTGGTGGCTGTGCTGCTGAGCCTGGAA
CTGCAGGAGCTGCCAGGTTATGGAAGGGCCAAGTGGAAATAAGCCCTTCGGCATTG
GGTACATGTGGTCGTACCATTAAATTGGCCTCATCTGGACTAGTGAATTCATCCTGGTGC
AGCAAATGACTATAGCTGGGCAGTGGTTACTTGTATTTAACAGAAGTAAAATGATCCTC
CTGATCATCCCACCTTCGTCTCTCCATTCTTCTTACCATCAAGGAACC GTTGTGA
AAGGGTCATTTAATCTCTGTGGTAGGGATTCCGAGAATCATTGTATGTACATGCAAACG
CACTGAAAGAACAGCAGCATGGTCATTGTCCAGGTACCTGTTCCGATGCTGCTACTGCTGTT
TCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATATACTACAAC TGCTATT
ATGGGACAGATTCTGTACATCAGCAAAGATGCATTCAAATCTGTCCAAGAACTCAAGTC
ACTTTACATCTATTAACTGCTTGGAGACTTCATAATTTCTAGGAAAGGTGTTAGGGTGT
GTTCACTGTTGGAGGACTCATGGCTTTAACTACAATGGCATTCCAGGTGTGGCAG
TCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTAGCCCAGTTTATCTGTGTTG
AAACTGTGCTGGATGCACCTTCCGTGTTGCTGTTGATCTGGAAACAAATGATGGATCGT
CAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCGTAAAAGGGAGCAACAAATTAA
ACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGAGGGAACAGAACTCCAGG
CCATTGTGAGA**TAG**ATACCCATTAGGTATCTGTACCTGGAAAACATTCTAAGAGCCA
TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTAAAAGACCTAA
TAAACCCATTCTCCTCAAAA

33/168

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCCCTGTTCTCCTTA
GAATAATTGTATGGGATTGTGATGCAGGAAAGCTAACAGGGAAAAAGAATATTCAATTCTGTG
TGGTGAAAATTTGAAAAAAAATTGCCTCTCAAACAAGGGTGTCAATTCTGATATTTAT
GAGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTGCTGGTGAC
TGGAGTACATTCAAACAAAGAAACGGCAAAGAAGAGATTAAAGGCCAAGTTCACTGTGCCTCA
GATCAACTGCGATGTCAAAGCCGAAAGATCATCGATCCTGAGTTCAATTGTGAAATGCCAGC
AGGATGCCAAGACCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGT
GTGTGGCGTGCCTACACAGTGGTGTGCTGATAATTCAAGGAGGGAAAATACTTGTGCGGAA
GGTTGCTGGACAGTCTGGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACC
ACGATGGAGAGAATCCTTATCGCTTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATC
AGCTCTACATACATCATCGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAGCCTA
TCAGAGGCCACCTATTCCAGGGACAACGTGCACAGCCGGTCACTCTGATGCAGCTCTGGCTGT
CACTGTAGCTGTGCCACCCCCACCCACCTGCCAAGGCCATCCCTCTGCTGCTTCTACCAC
CAGCATCCCCAGACCACAATCAGTGGCCACAGGAGCCAGGAGATGGATCTGGTCCACTGC
CACCTACACAAGCAGCCAAACAGGCCAGAGCTGATCCAGGTATCAAAGGCAAGATCCTTC
AGGAGCTGCCTTCCAGAAACCTGTTGGAGCAGTGTGAGCGGATGTCAGCTGGGACTTGTCCAAAAGAAGA
ATTGAGCACACAGTCTTGGAGCAGTATCCCTGGAGATCCAAACTGCAAATTGACTTGT
GTTTTAATTGATGGGAGCACCAGCATTGGCAAACGGCGATTCCGAATCCAGAAGCAGCTCCT
GGCTGATGTTGCCAAGCTCTGACATTGGCCCTGCCGGTCCACTGATGGGTGTTGTCAGTA
TGGAGACAACCTGCTACTCACTTAACCTCAAGACACACAGAATTCTCGAGATCTGAAGAC
AGCCATAGAGAAAATTACTCAGAGAGGAGACTTCTAATGTAGGTCGGGCATCTCCTTGT
GACCAAGAACTTCTTCCAAAGCCAATGGAAACAGAAGCGGGCTCCAAATGTGGTGGTGGT
GATGGTGGATGGCTGGCCCACGGACAAAGTGGAGGGCTCAAGACTTGCAGAGAGTCAGG
AATCAACATTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAGAAGCAGTATGTGGTGG
GCCCAACTTGCACAAACAAGGCCGTGTGCAGAACAAACGGCTCTACTCGCTCACGTGCAGAG
CTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCGGGCTGCGACACTGACCGCCT
GGCCTGCAGCAAGACCTGCTGAACTCGGCTGACATTGGCTCGTCATCGACGGCTCCAGCAG
TGTGGGGACGGCAACTCCGCACCGTCCAGTTGTGACCAACCTCACAAAGAGTTGA
GATTCGCACACGGACACCGCGCATGGGGCGTGCAGTACACCTACGAACAGCGGCTGGAGTT
TGGGTTGACAAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGCTACTG
GAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTGCCTGGAGCAGCTTCAAGAAGTC
CAAGCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCAGGGAGGTCTACGACGACGTCCG
GATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCGATAGGCCTGCTGGC
TGCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCAACTCCTCTTGT
CGAGTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAA
CTCACAGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGTGCTTACTA
ACGTGTTGGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTGGCAGGGCA
TGGAGAAACAAATGTCTGTTATTATTCTTGCCATCATGCTTTCATATTCCAAAACCTGG
ACTTACAAAGATGATCACAAACGTATAGAATGAGCAGGCTACATCATGTTGAGGGTGT
GGAGATTTACATTGACAATTGTTCAAATAATGTTGGAATACAGTGCAGCCCTAC
GACAGGCTACGTAGAGCTTGTGAGATTAAAGTTGTTATTCTGATTGAACTCTGTAA
CCCTCAGCAAGTTCATTTGCTGACAAATGTAGGAATTGCTGAATTAAATGTTAGAAGG
ATGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCGTGTGGAAAAGCTATCAAGGAA
GAAATTGCCAAACCATGTCTTTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGT
TAATTAAGCATGGAATACAGAAAACAACAAAAACTTAAGCTTAATTCATCTGGAATTCCA
CAGTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTCCCGCTGGCTGCTATCAC
GTGGTGCTCTCCGACTACTCACCCGAGTGTAAAGAACCTCGGCTCGGTGCTCTGAGCTG
CTGTGG**ATG**GCCTCGGCTCTGGACTGTCCTCCGAGTAGGATGTCACTGAGATCCCTCAA
TGGAGCCTCCTGCTGTCACTCCTGAGTTCTTGTGATGTGGTACCTCAGCCTCCCCAC
TACAATGTGATAGAACCGGTGAACGGATGTACTCTATGAGTATGAGCCGATTTACAGACAA
GACTTCACTTACACTTCGAGAGCATTCAAACGTCTCATCAAATCCATTCTGGTCATT
CTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTGGGTGAA
AAAAAGTCTTGGTGGGATATGAGGTTCTACATTTCTTATTAGGCCAAGAGGGCTGAAAAG
GAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTATGGTACATAATCCGA
CAAGATTTTAGACACATATAATAACCTGACCTGAAAACCATTATGGCATTCAAGGTGGTA
ACTGAGTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATGTTTCAATAACT
GGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGAGAAGTTTACAGGTTATCCT
CTAATTGATAATTATTCCATAGAGGATTTACCAAAAAACCCATATTCTTACCAAGGAGTAT
CCTTCAGGTGTTCCCTCCATACTGCAGTGGTTGGTTATATAATGTCCAGAGATTGGT
CCAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGG
ATCTGTTGAATTATTAAGTGAACATTCAATTCCAGAACAGACACAAATCTTCTTCTA
TATAGAATCCATTGGATGTCTGTCACGTGAGACGTGTGATTGCAGCCATGGCTTCTC
AAGGAGATCATCACTTTGGCAGGTATGCTAAGGAACACCACATGCCATTAT**TAA**CTTCAC
ATTCTACAAAAAGCCTAGAAGGACAGGATACCTGTTGGAAAGTGTAAATAAGTAGGTACTG
TGGAAAATTGATGGGGAGGTCAAGTGTGCTGGCTTACACTGAAACTCATGAAAAACCC
GACTGGAGACTGGAGGGTTACACTGTTGATTATTAGTCAGGCCCTCAAAGATGATATGTGG
AGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGGACCAAACAATTG
GACATGTCATTCTGTAGACTAGAATTCTTAAAGGGTGTACTGAGTTATAAGCTCACTAGG
CTGTAAAAACAAACATGTAGAGTTATTGAAACAATGTAGTCACGTGAAGGTTTGT
GTATATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTGTCAAAAAAACTTCTCA
CTGAAGTTACTGAACAAAATTACCTGTTGGTCATTATAAAAGTACTTCAAGATGTT
GCAGTATTTCACAGTTATTATTTAAATTACTCAACTTGTGTTTAAATGTTTGAC
GATTCAATACAAGATAAAAAGGATAGTGAATCATTCTTACATGCAAACATTCCAGTTAC
TTAAGTGTAGTTATTGATAACATCACTCCATTAAATGTAAAGTCAGGTCAATTATTG
ATATCAGTAATCTCTGGACTTGTAAATATTTACTGTGGTAATATAGAGAAGAATTAAAG
CAAGAAAATCTGAAAAA

FIGURE 37

39/168

FIGURE 39

GGTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCGTGATTATTAACGTGGCTTA
ATCTGAAGGTTCTCAGTCAAATTCTTGTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTT
GGCTGGTTGGGCCCTGTAGCTGACAAGGTGCCAGGGAGAATGCAGCACACTGCTCGAGAATGAAGGC
GC TTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGCAACCTGCACTTCTGTATT
CAGAACTCTGTAAAGGTGCCTCCACTACGGCTGACCAAAGATAGGAAGAGGGCGCTACAAGATGGCTGTCCAG
ACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCTTAATGACAG
ACGAGCCTGGCCTAGACAACCCCTGCCTACGTGTCTGGCAGAGGACGGCAGCCAGCAATCAGCCCAGTGGACT
CTGGCCGGAGCAACCGAACTAGGGCACGGCCTTGAGAGATCCACTATTAGAACGAGATCATTAAAAAAATAA
ATCGAGCTTGAGTGTCTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAA
ATTCTGAAAACACCACTGCCCTGAAGTCTTCCAAGGTTGATACCACCTGATTCCAGATGGTGAATTACCAAGCA
TCAAGATCAATCGAGTAGATCCAGTGAAGGCCTCTCTATTAGGCTGGTGGAGGTAGCAGAACCCCCACTGGTCC
ATATCATTATCCAACACATTTATCGTATGGGTGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC
TAAAGGTCAACGGGATGGACATCAGCAATGTCCCTACAACACTACGCTGTGCGTCTCTGCCAGGCCCTGCCAGG
TGCTGTGGCTGACTGTGATGGTGAACAGAAGTCCGCAGCAGGAACAATGGACAGGGCCGGATGCCACAGAC
CCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCCCGAGGAGCAGCTTGAATAAAACTGGTGCAGCA
AGGTGGATGAGCCTGGGTTTATCTCAATGTGCTGGATGGCGGTGTGGCATATGACATGGTCAGCTTGAGG
AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCAGCCAGAAAAGTGCAGCTCATCTGATT
AGGCCAGTGAAGAGCGTGTTCACCTCGTGTCCCGCCAGGTTGGCAGCGGAGCCCTGACATCTTCAGGAAG
CCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGAGCAACACTCCAAAGCCCCTCCATCCTA
CAATTACTTGTGATGAGAAGGTGAAATATCCAAAAGACCCGGTGAATCTCGGATGACCGTGCAGGG
GAGCATCACATAGAGAATGGGATTGCTATCTATGTGATCAGTGTGAGGCCGGAGGAGTCATAAGCAGAGATG
GAAGAATAAAACAGGTGACATTGTTGAAATGTGGATGGGTCGAACGTGACAGAGGTGAGCCGGAGTGAGGCAG
TGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGAAGTCAAAGAGTATGAGCCCCAGGAAAG
ACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACACATGGCCCCACCCAGTGAUTGGTCCCCATCTGGTCA
TGTGGCTGGAATTACCACGGTGTCTTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTCTGG
GCTTCTGCATTGAGGAGGTATGAAGAATACAATGGAAACAAACCTTTTCTCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTATTTCTTGTGTCATGGTAGAAGTACATCAGGAA
TGATACATGCTGCTGGCAAGACTGCTGAAAGAAACTAAAGGAAGAATTACTCTAACTATTGTTCTGGCTG
GCACTTTTTATAGAATCAATGATGGTCAGAGGAAACAGAAAATCAAATAGGCTAAGAAGTTGAAACACT
ATATTTATCTGTGAGTTTATTTAAAGAAAGAACATATTGTAAGGAAACTGTCAGGAAAGTATGATCATCTAA
TGAAAGCCAGTTAACCTCAGAAAATGATTCAACAAAATTAAGGAAACTGTCAGGAAAGTATGATCATCTAA
TTCTCATTACTCTACAACATTGTTATTTTCTATTCAATAAAAGCCTAAAACAACAAATGATTGATT
TGTATAACCCACTGAATTCAAGCTGATTAAATTAAATGGTATATGCTGAGCTGCTGCCAAGGGTACATTAT
GCCATTTAATTACAGCTAAATATTAAAATGCTGAGAACGTTGCTTCAACACAAGAAT
AAATATTTCAGAAGTTAAA

41/168

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAA
GCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAAC**ATG**GGCTT
CAACCTGACTTTCCACCTTCCACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCCTGAC
AGTGGTTGGGTGGGCCACCAGTAACTACTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAA
GGAGTTCATGGCTAATTCCATAAGACCCCTCATTGGGGAAAGGGAAAAACTCTGACTAATGA
AGCATCCACGAAGAAGGTAGAACCTGACAACACTGTCCTCTGTCTCCTTACCTCAGAGGCCA
GAGCAAGCTCATTTCAAACACAGATCTCACCTTGAAGAGGGTACAGGCAGAAAATCCCAAAGT
GTCCAGAGGCCGGTATGCCCTCAGGAATGTAAGCTTACAGAGGGTGCCTCCTCGTTCC
CCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTCCTGCAGAG
GCAGCAGCTGGATTATGGCATCTACGTACCCACAGGCTGAAGGTTAAAGCTTAAATCGAGC
CAAACCTTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGACTGCTTATATT
CCACGATGTGGACCTGGTACCCGAGAATGACTTAAACCTTACAAGTGTGAGGAGCATCCAA
GCATCTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACAGTGGATATTGGGGGG
TGTTACTGCCCTAACGAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGG
ATGGGGAGGCAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCG
GCCCTGCCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGCAATGAGGT
GAACGCAGAACGGATGAAGACTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGTTGAG
TAGTTGTTCTTATAAAATTAGTATCTGTGGAACACAATCCTTATATACACATCACAGTGG
TTTCTGGTTGGTGCA**TGA**CCCTGGATCTTGGTATGTTGGAAGAACTGATTCTTGT
GCAATAATTGGCCTAGAGACTTCAAATAGTAGCACACATTAAGAACCTGTTACAGCTCATT
GTTGAGCTGAATTTCCTTTGTATTTCTTAGCAGAGCTCCTGGTATGTAGAGTATAAA
ACAGTTGTAACAAGACAGCTTCTTAGTCATTTGATCATGAGGGTAAATATTGTAATATGG
ATACTTGAGGACTTATATAAAAGGATGACTCAAAGGATAAAATGAACGCTATTGAGGACT
CTGGTTGAAGGAGATTATTAAATTGAAGTAATATATTATGGATAAAAGGCCACAGGAAA
TAAGACTGCTGAATGTCTGAGAGAACAGAGTTGTTCTCGCCAAGGTAGAAAGGTACGAAGA
TACAATACTGTTATTCAATTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGGTGAGAAG
GCGTCCACAAAAGAGGGAGAAAAGGCAGAACAGTGAACCTGGGAATGAAGAG
GTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGA
TAGCCTTCAGGGGAGGACCTGCCAGGTATGCCCTCCAGTGATGCCACCAAGAGAACATTC
TCTATTAGTTTAAAGAGTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGT
TTACAAGTTACATATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTG
AAAAAGCAAAA

43/168

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATGG**CACTGAGCTCCAGATCTG
GGCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTCCCACA
ACAGACGGGACAAC TGCAAGAGCTGCAACCCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGAT
GCCCATGTTCCAGAGGCAGAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGC GG
CTGCTGTCA CGATCAAAGTGTGGATGTGCTGCAAGACG**TAGAAC**CTACCTGCCCTGCC CCCCC
GTCCCCCTCCCTCCTTATT TATT CCTGCTGCC CAGAACATAGGTCTTGGAAATAAAATGGCTG
GTTCTTTGTTTCCA AAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA AAAAAAAAAA
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

45/168

FIGURE 45

GTGGCTCATTCACTGGCTGACTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCAC
CCTCATCTATACCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGT
CGGTTCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTAT
TGTCTGGACCTTCAACACAACCCCTTGTCAACCATAAGCCAGAAGGGGGCACTATCATAGT
GACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAG
CAAACGTGAAGAAGAATGACTCAGGGATCTACTATGTGGGATATACAGCTCATCACTCCAGCA
GCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCAT
GGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGG
GGAAGAGGATGTGATTTACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGG
GTCCATCCTCCCCATCTCCTGGAGATGGGAGAAAGTGATATGACCTCATCTGCCTGCCAG
GAACCTGTCAGCAGAAACTCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGC
TGATGACCCAGATTCCCTCATGGTCCTCCTGTGTCTCCTGTTGGTGCCTCCTGCTCAGTCT
CTTGACTGGGCTATTCTTGTTCTGAAGAGAGAGACAAGAAGAGTACATTGAAGA
GAAGAAGAGAGTGGACATTGTCGGAAACTCTAACATATGCCCTATTCTGGAGAGAACAC
AGAGTACGACACAATCCCTCACACTAACATAGAACATCCTAAAGGAAGATCCAGCAAATACGGT
TTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCACTGCTCACGATGCCAGA
CACACCAAGGCTATTGCCTATGAGAATGTTATCTAGACAGCAGTGCACCTCCCTAAGTCTCT
GCTCA

47/168

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACC**ATG**ACCTGCTGCGAAGGATGGACATCCTGCAAT
GGATTCA~~G~~CCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATA~~CC~~TCTAATTGTC
AGCTTAGTTGAGGAAGACCAATTCTCAAAACCCATCTCTGCTTGAGTGGTGGTCCC
GGAATTATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAA
AGAGCGTGCTGCAACAA~~C~~AGAAC~~T~~GGAA~~T~~GT~~T~~TCTTCATCATTTCAGTGTGATCACAGTC
ATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTTAAAAGGT~~C~~CTCTCATGTGT
AATTCTCCAAGCAACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATT~~C~~
CCAGAATCCTTCAACTTG~~C~~AGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAAT
AAACCCACCAGTAACGACACC~~A~~GGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCT
GAAGAAAAACAAACATAGGCTTATCCACTTCTCAGTATT~~T~~TAGGTCTATTGCTTGGAAATT
CTGGAGGT~~C~~CTGTTGGCTCAGTCAGATAGTCATCGGTT~~C~~TTGGCTGTGTGGAGTC
TCTAAGCGAAGAAGTCAAATTGTG**TAG**TTAATGGAAATAAAATGTAAGTATCAGTAGTTGA
AAAAAAAAAA

49/168

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGT
GAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGA
GCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAA
CTGTGAC**ATG**GAGAGAGTGACCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGC
CAATGACCCATTGCCAATAAAGACGATCCCTRACTATGACTGGAAAAACCTGCAGCTGAG
CGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGATCGGGCAGTTCTGAGTGGCAAATG
CAAATACAAGAGCAGCCAGAACGACAGCACAGTCCTGTACCTGAGAAGGCCATCCACTCATCAC
TCCAGGCTCTGCCACTACTTGCT**TGA**GCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAAC
ACTGGCCCCAGCACCTCCTCCCCCTGGGAGGCCTATCCTCAAGGAAGGACTTCTCTCCAAGG
GCAGGCTGTTAGGCCCTTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACC
CCCTCA

51/168

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAAGGCTGCAGACCCAGAGGGAG
GGAGGACAGGGAGTCGGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGC
AAGGAGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAG
TTCCAGGGGCCCTGGCCTGCCTCCTGCTGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCC
CTGCAGAGCGGAGAGGAAAGCACTGGACAAATATTGGGAGGCCCTGGACATGCCCTGGGA
GACGCCCTGAGCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCT
AAAGTCAGTGAGGCCCTGGCCAAGGGACAGAGAACAGCAGTTGGCACTGGAGTCAGGCAGGTT
CCAGGCTTGGCGCAGCAGATGCTTGGCAACAGGGTCGGGAAGCAGGCCATGCTCTGGGA
AACACTGGGCACGAGATTGGCAGACAGGCAGAAGATGTCATTGCACACGGAGCAGATGCTGTC
CGCGGCTCCTGGCAGGGGTGCCTGCCACAGTGGCTTGGAAACTCTGGAGGCCATGGC
ATCTTGGCTCTCAAGGTGGCCTGGAGGCCAGGGCAATCCTGGAGGTCTGGGACT
CCGTGGGTCCACGGATAACCCCGAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCT
CCCTGGGTCAAGGAGGCCAATGGAGGGCACCAAACACTTGGCACCAACACTCAGGGAGCTGTG
GCCAGCCTGGCTATGGTCAGTGAGAGCCAGCAACCAGAACAGGGTGACGAATCCCCCA
CCATCTGGCTCAGGTGGAGGCTCCAGCAACTCTGGGGAGGCAGCAGCTCAGTCGGCAGC
AGTGGCAGTGGCAGCAATGGTGACAACAATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGC
AGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGCAGCAGTGGTGGCAGCAGTGGCAAC
AGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAATCCTCCTGGGATCCAGCACCGGCTCCTCC
TCCGGCAACCACGGTGGGAGCGGGAGGAAATGGACATAAACCGGGTGTAAAAGCCAGGG
AATGAAGCCCAGGGAGCGGGAAATCTGGATTAGGGCTTCAGAGGACAGGGAGTTCCAGC
AACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTGGAGGCTCTGGAGACAATTATCGG
GGCAAGGGTCAGCTGGGCAGTGGAGGGTGAAGCTGGTGGAGTCAATACTGTGAAC
TCTGAGACGTCTGGATGTTAACACTTGCACACTTCTGGAGAATTAAATCCAAGCTG
GGTTCATCAACTGGATGCCATAAACAAAGGACCAGAGAACAGCTCTGCATCCCGTGACCTCCA
GACAAGGAGCCACCAGATTGGATGGAGGCCACACTCCCTCTTAAACACCACCCCTCTCA
TCACTAATCTCAGCCCTGGCTGAAATAACCTTAGCTGCCCAACAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

53/168

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGGA**TGTCGCTGCTGAGCCTGCCCT**
GGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCT
GGCTACTCGCCCGCATCCTGGCTGGACCTATGCCTTCTATAACAACGTGCCGCCGGCTCCAGT
GTTTCCCACAGCCCCAAAACGGAACGGTTTGCGGTCACCTGGGCTGATCACTCCTACAG
AGGAGGGCTGAAGGACTCGACCCAGATGTCGCCACCTATTCCCAGGGCTTACGGTATGGC
TGGGTCCCACATCCCCTCATCGTTTATGCCACCCTGACACCATCCGGTCTATACCAATG
CCTCAGCTGCCATTGCACCCAAGGATAATCTCTCATCAGGTTCTGAAGCCTGGCTGGAG
AAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCT
TCCATTCAACATCCTGAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGC
ACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCC
TCATGACCTTGGACAGTCTACAGAAATGCATCTCAGCTTGACAGCCATTGTCAGGAGAGGC
CCAGTGAATATATTGCCACCATCTGGAGCTCAGTGCCCTGTAGAGAAAAGAACGCCAGCATA
TCCTCCAGCACATGGACTTCTGTATTACCTCTCCATGACGGCGGCCTCCACAGGGCCT
GCCGCCTGGTGCATGACTCACAGACGCTGTCATCCGGAGCGGGCGTCGCACCCCTCCCCACTC
AGGGTATTGATGATTTTCAAAGACAAAGCCAAGTCCAAGACTTGGATTTCATTGATGTGC
TTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGATATAAGAGCAGAGGCTG
ACACCTTCATTTGGAGGCCATGACACCACGGCAGTGGCCTCTCCTGGGCTGTACAACC
TTGCGAGGCACCCAGAATACCAGGAGCGCTGCCACAGGAGGTGCAAGAGCTTCTGAAGGACC
GCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCCTTCCTGACCATGTGCGTGA
AGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTCATCTCCGATGCTGCACCCAGGACATTG
TTCTCCCAGATGGCGAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAAGGGTCC
ATCACAACCCAACGTGTGGCCGGATCCTGAGGTCTACGACCCCTCCGCTTGACCCAGAGA
ACAGCAAGGGAGGTACCTCTGGCTTTATTCCCTTCTCCGAGGGCCCAGGAACGTGAC
GGCAGGCAGTTCGCCATGGCGGAGATGAAAGTGGCCTGGCGTTGATGCTGCTGCACCTCCGGT
TCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGCGCGCCAGGGCGGGC
TTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGAG**TGA**CTTCTGACCCATCCACCTGTT
TTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

55/168

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTC**ATGG**GACCAGTGAAACAGCTGAAGCGAATGTTT
GAGCCTACTCGTTGATTGCAACTATCATGGTGCTGTGTGTTGCACTTACCCGTGTTCT
GCCTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGGCATTG
ACGTGGTACAGCCTTCCTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTG
TGTCTTGCA**TAA**TTCATGCCAGTTTATGAAGCTTGGAGGCACATGGACAGAACGCTGGT
GGACAGTTTGTAACTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTATCTGCAGC
AATGTGTTGCTTGATTGCAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAA
CCTGAATGTCAGTAGCACAGGATGAGAAGTGGTTCTGTATCTGTGGAGTGGAAATCTTCCTC
ATGTACCTGTTCTCTGGATGTTGCCACTGAATTCCCATGAATAACAAACCTATTCAAGC
AACAGCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

57/168

FIGURE 57

CGGCTGAGGCTGAGCCGAATCGGCTGAGGGCAGTGGAGCACCCAGCAGGCCAAC**ATG**CTCTGTCTGTGC
CTGTACGTGCCGGTCATCGGGGAAGGCCAGACCGAGTTCCAGTACTTGAAGTCGAAGGGCTCCCTGCCAGCTG
AAGTCCATTCAAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCCACCTACGCCAGTGGAAAGCAGAAAATT
GTACAAGCTGGAGATAAGGACCTTGATGGCAGCTAGACTTGAAGAATTGTCCATTATCTCAAGATCATGAG
AAGAAGCTGAGGCTGGTGTAAAGATTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAG
TCCCTGCCGGACTTGGGAGTCAAAGATATCTAACAGCAGGGAGAAAAATTCTCAAGAGCATGGATAAAAACGGC
ACGATGACCATCGACTGGAACGAAGTGGAGAGACTACCACCTCCTCCACCCGTGGAAAACATCCCCGAGATCATC
CTCTACTGGAAGCATTCCACGATCTTGATGTGGTGAGAATCTAACGGTCCGGATGAGTTACAGTGGAGGAG
AGGCAGACGGGATGTGGTGGAGACACCTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCC
CCCCCTGGACACGGCTCAAGGTGCTATGCAGGTCCATGCCTCCCGCAGCAAACACATGGCATCGTTGGGCTTC
ACTCAGATGATTGAGAAGGAGGGCCAGGTCACTCTGGCGGGCAATGGCATCAACGCTCCCTCAAATTGGCCCC
GAATCAGGCATCAAATTCATGGCCTATGAGCAGATCAAGCGCCTTGTGGTAGTGGACAGGAGACTCTGAGGATT
CACGAGAGGGCTGTGGCAGGGCTTGGCAGGGGCATCGGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAG
ACCCGGATGGCGCTCGGAAGACAGGCCAGTACTCAGGAATGCTGGACTCGGCCAGGAGGATCCTGGCCAGAGAG
GGGGTGGCCCTTCATACAAAGGCTATGTCCTTGGCAGGGCAGCTGGCAGGAGGATCCTGGCCAGAGAG
GTCTACGAGACGCTCAAGAATGCCCTGGCTCAGCAGTATGCACTGAACAGCGCAGCCGGCTGGCTGGCT
CTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCAGCTACCCCTGGCCCTAGTCAGGACCCGGATG
CAGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTTCAAACATATCCTGGGACCGAG
GGGGCCTCGGGCTGTACAGGGGCTGGCCCCAACCTCATGAAGGTCACTCCAGCTGTGAGCATCAGCTACGTG
GTCTACGAGAACCTGAAGATCACCTGGCGTGCAGTCGCG**TGA**CGGGGGAGGGCCGCCAGTGGACTCG
CTGATCCTGGCCGCAGCCTGGGTGTGCAGCCATCTCATTGTGAATGTGCCAACACTAACGCTGCTCGAGCC
AAAGCTGTAAAACCTAGACGCAACCGCAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTGTCTGCTGACC
CCAGCAGACCCCTCTGTGTTGGTCCAGCGAAGACCAACAGGCACTTCTTAAGGGTCCAGGGTCAGCAGGCTCCGGGCT
CACATGTGTAAGGACAGGACATTCTGCAGTGCCTGCCAATAGTGAGCTGGAGGCTGGAGGCGCTTAGTTC
TTCCATTTCACCCCTGCAGCCAGCTGTGGCCACGSACCTGCCCTCTGGCTGCCGTGCATCTCCCTGTGCCCT
CTTGCTGCCTGCCGTGCTGCTGAGGTAAGGTGGAGGGCTACAGCCCACATCCCACCCCTCGTCAAATCCC
ATAATCCATGATGAAAGGTGAGGTACCGTGGCCTCCAGGCCTGACTTCCAAACCTACAGCATTGACGCCAACCT
GGCTGTGAAGGAAGAGGAAGGGATCTGGCCTGTGGCTACTGGCATCTGAGCCCTGTGATGGCTGGGCTCATGGT
GGCATGCTGGAGTGCAGGGGCTCGGGCTGCCCTGGCTGACAGAAGGCAAGTGTGGCTGGGCTCATGGT
CTCTGAGCTGGCCTGGACCCCTGTGAGGATGGGCCCCACCTCAGAACCAAACACTACTGTCCCCACTGTGGCATGAG
GGCAGTGGAGCACCATGTTGAGGGCGAAGGGCAAGCGTTGTGTTCTGGGAGGGAGGAAAGGTGTTGG
AGGCCTTAATTATGGACTGTTGGAAAAGGGTTTGTCCAGAAGGACAAGCCGGACAATGAGCAGCTCTGTG
TTCCAGAGGAAGACGAGGGAGCAGGAGCTGGCTACTGTCAGAGTCTGTTCTGACGCCCTGGGGTTCTGTG
CAACCCAGCAGGGCGCAGCGGACGACCCACATTCACTGTGTCAGTGGAACTATTATTTGTA
TTTATTGAAACAGAGTTATGCTTAACATTTTATAGATTGTTAATTAAATAGCTGTGTCATTCAAGTTCAT
TTTTATTCAATTATGTTCATGGTATTGTCACCTCCCAAGGCCGGCAGTGGGATGGGAGGAGGAGGAGAA
GGGGGCGCTGGGCCGCTGCAGTCACATCTGTCAGAGAAATTCTTTGGACTGGAGGAGAAAAGCCGGCAG
AAGGCAGCAGCCCTGGCTCTTCCCTTGGCAGGGTGGGAAGGGCTGCCAGGCCCTAGGATTCAGGGTT
GACTGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTGAGGTTGGAAATCCAGTTATTCTGCCGTGCGA
GGGTTTCTTATTTCACTCTTCTGAATGTCAGGCAGTGGAGGTGCCCTCTCACTGTGAAATTGTGGTGGCGGG
GGCTGGAGGAAGGGTGGGGCTGGCTCCGTCCCTCCAGGCCCTCTGCTGCCCTGCTTAACAATGCCGGCCAA
CTGGCGACCTCACGGTTGCACTCCATTCCACAGAATGACTGAGGAAATCTCAATAGGATGCAAGAGATC
AATGCAAAATTGTTATATGAAACATATAACTGAGTCGTCAAAAGCAAAATTAAAGAAAGAATTGGACGTTAGA
AGTTGTCATTAAAGCAGCCTCTAATAAAAGTTGTTCAAAGCTGAAAAAAGAAAAAAGAAAAAAGAAAA
AAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA

59/168

FIGURE 59

GGAAGGCAGCGGAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTCCCCAGCC **ATG**
GCTTCCCTGGGGAGAT CCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCA
ATTGCACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCC
TCAGCTGGAACATTGGGAGGATGGAATCCTGAGCTGCAC TTTGAACCTGACATCAAAC TT
TCTGATATCGT GATA CAATGGCTGAAGGAAGGTGTTAGGCTGGTCCATGAGTTCAAAGAA
GGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTT CAGAGGCCGGACAGCAGTGTTGCTGAT
CAAGTGATAGTGGCAATGCCTCTTGC GGCTGAAAAACGTGCAACTCACAGATGCTGGCACC
TACAAATGTTATATCATCACTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGA
GCCTCAGCATGCCGGAA GTGAA ATG TGGA CTATAATGCCAGCTCAGAGACCTTGC GG TGAG
GCTCCCCGATGGTCCCCCAGCCCACAGTGGCTGGCATCCAAAGTTGACCAGGGAGCCAAC
TTCTCGGAAGTCTCCAATACCAGCTTGAGCTGAACTCTGAGAATGTGACC ATGAAGGTTGTG
TCTGTGCTCTACAATGTTACGATCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCC
AAAGCAACAGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCCGGAGTCACCTACAGCTG
CTAAACTCAAAGGCTTCTGTGTCTCTTCTTGC CATCAGCTGGCACTTCTGCCT
CTCAGCCCTTACCTGATGCTAAAAT**TAA**TGTGCCTGGCCACAAAAAGCATGCAAAGTCATTG
TTACAACAGGGATCTACAGAACTATTCACCACCAGATATGACCTAGTTATATTCGGGA
GGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCAAGAAACAAAAAGAAGCC
AAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTCAAAGACATATTAGAAGTTGGG
AAAATAATT CATGTGAACTAGACAAGTGTGTTAAGAGTGATAAGTAAAATGCACGTGGAGAC
AGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGAGTGAGAGGACAGGAT
AGTGCATGTTCTTGTCTGAATT TTAGTTATGTGCTGTAATGTTGCTCTGAGGAAGCC
CCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCACAAATTAGCTGTAGTATGTAC
CCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGCGGCTGCATT TAGTAATGGG
TCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTGGCTCTTCCAACTGACAAA
TGCCAAAGTTGAGAAAAATGATCATAATT TAGCATAAACAGAGCAGTCGGGACACCGATTT
TATAAATAAAACTGAGCACCTTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

61/168

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACA
AGCACCAAGGAGCCCCTCCGGTAGCTACTACCCTGGACCCCCAATAGTGGAGGGCAGTATGG
TAGTGGGCTACCCCTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTATGGACCACC
AGCTGGTGGAGGGCCCTATGGACACCCAATCCTGGATGTTCCCTCTGGAACCTCAGGAGG
ACCATATGGCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGC
CCAGCAGCCTGGCTTATGGACAGGGTGGCGCCCTCCAATGTGGATCCTGAGGCCTACTC
CTGGTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGC
CCTGGTCAACTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTT
TGACAAGACCAAGTCAGGCCGCATCGATGTCTACGGCTTCAGCCCTGTGGAAATTCACTCCA
GCAGTGGAAGAACCTCTCCAGCAGTATGACCGGGACCGCTCGGCTCCATTAGCTACACAGA
GCTGCAGCAAGCTCTGCCAAATGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGT
CTCCCGCTACTGCCACGCTCTGCCAATCCTGCCATGCAGCTGACCGCTTCATCCAGGTGTG
CACCCAGCTGCAGGTGCTGACAGAGGCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACAT
CCGGCTCAGCTCGAGGACTTCGTCACCATGACAGCTCTCGGATGCTATGACCCAACCATCT
GTGGAGAGTGGAGTGCACCAGGGACCTTCCTGGCTTCTAGAGTGGAGAGAAGTATGTGGACA
TCTCTTCTTCTGTCCCTCTAGAAGAACATTCTCCCTGCTGATGCAACACTGTTCCAAA
AGAGGGTGGAGAGTCCTGCATCATGCCACCAAATAGTGAGGACCGGGCTGAGGCCACACAG
ATAGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCCTGATGCCATGAGCAGTTGAGTG
GCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCA
CCCTGATGCCAGTGGTGAGTGTCTGCCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAG
GCCATCCTGTCAAACGAGCCCATTCTCAAAGTGGAACTGACCAAGCATGAGAGAGATCT
GTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAATCCTGTGTTAACTCTAGC
TGCCTGGGCTGCCCTGCTCAGACAAATCTGCTCCCTGGCATCTTGGCAGGCTCTGCC
CCCTGCAGCTGGACCCCTCACTTGCCTGCCATGCTCTGCTCGGCTTCAGTCTCCAGGAGACA
GTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCACTTTTCAATTGGGCCAAAAG
TCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

63/168

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATCTCCCAGGCCTTTGCCCGGAGCAGCTGGTGTGGAGGAGAAAGTTCCAAAACCTCGGGACCAACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCTCTAACACTGAACATCCGCA GCCCGCTCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCTGAAGCTCAGCGTGCCTCCATCAGATGGCTTCCCACCTGCAGGAGGTTCTGCAGTCAGAGGTGGCCTCCATCGTGGGGCT GCCTGCCATGGATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCGGCTGAGGA CCGCCTGGGGGAAGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCTGCCACAGGCCTCTCACCTGAGGCTTC GGGCAGTGGCCCTTGCCTGGGAGTCTTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTC ACTCCTCCACCAAGGACTCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGGCCGGGG AAAAATCCTTCCAACGCCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCCTGAATCCCAGTGTGTCCTGGGAGGTTGGAGGCCCTGGACTGGTTGGGAACGAG GCCCATGCCACACCCCTGAGGAATCTGGGTATCAATAATCAACCCCCAGGTACCACTGGGG AAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTG GGGGAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTCACTATACCCAGGTATCAA TAACCCATTCCTCCCTGGAGTTCTCCGCCCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTT CCCTAACCTCCAAGCCCTAGGTGCAAGTGGGCTAGAGCACGATAAGGGAAACCCAACATT GGGAGTTAGAGTCCTGCTCCGCCCTTGCTGTGGCTCAATCCAGGCCCTGTTAACATGTT CCAGCACTATCCCCACTTTCAAGTGCCTCCCTGCTCATCTCCAATAAAATAAAAGCACTT ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCACTT AAAAAAAAAAAAAAAAAAAAAAA

65/168

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGG
TCTGGGCTGCCCCTTGTCTCCTCTTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGT
ATGACTTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCTCCTATGAGTCCAGCTTC
CTGGAATTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTC
CACCATGCAAGATCTAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCC
TTCTTGGCCCGGGCTTTGGCCGGGATGCAGGAGGCAGGCCCGACCCTGTCTTCAGCAG
GCCCCCACCCCTCCTGAGTGGCAATAAATAAAATTGGTATGCTG

67/168

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGG
GCCAGGTGCCCGTCGCAGGTGCCCTGCCGGAGATGCGTAGGAGGGCGAGCGCGAGAAG
CCCCCTCCTCGCGCTGCCAACCCGCCACCCAGCCC**ATG**GCGAACCCGGGCTGGGCTGCTT
CTGGCGCTGGGCCTGCCGTTCTGCTGGCCGCTGGGGCCGAGCCTGGGGCAAATACAGACC
ACTTCTGCAAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAAC
CTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTGGCTGCCTGCTCCTG
GCTGTGGGCTGGCACTGTTGGTGCAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTAC
CGGCCAGTAGCGAGGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAG
GAGACGGTGCAGGGCTGCCATC**TAG**GTCCCTCTGCATCTGTCTCCCTATTGC
TGTGTGACCTTGGGAAAGGCAGTGCCCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAG
CAGGGAAAGGTAATTCAAAGACTCTGCCCTGAGGTCAAGAGAGGGATGGGCTATTCACTT
TTATATATTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAA

69/168

FIGURE 69

GCCAGGAATAACTAGAGAGGAACA**ATGGGGTTATT**CAGAGGTTTGTTCCTCTTAGTTCTGTGCCTGCTGCAC
 CAGTCAAATACTTCCTCATTAAGCTGAATAATAATGGCTTGAAGATATTGTCAATTGTTATAGATCCTAGTGTG
 CCAGAAGATGAAAAAATAATTGAACAAATAAGGGATATGGTGAACAGCTTCTACGTACCTGTTGAAGCCACA
 GAAAAAAGATTTTTCAAAAATGTATCTATATTAAATTCTTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGG
 CCAAAACATGAAAACCATAAACATGCTGATGTTATAGTTGCACCCACCTACACTGCCAGGTAGAGATGAACCATAAC
 ACCAAGCAGTTCACAGAATGTGGAGAGAAAAGCGAATACATTCACTCACCCCTGACCTTCTACTTGGAAAAAAA
 CAAAATGAATATGGACCACCAAGSCAAACTGTTGTCCATGAGTGGCTCACCTCCGGTGGGAGTGTGATGAG
 TACAATGAAGATCAGCCTTCTACCGTCTAAGTCAAAAAAATCGAACAGCAACAGGTGTTCCGCAGGTATCTCT
 GGTAGAAATAGAGTTATAAGTGTCAAGGAGGCAGCTGTCTAGTAGACCATGAGAATTGATTCTACAAACAAAA
 CTGTATGGAAAAGATTGTCAATTCTTCCTGATAAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAAGT
 ATTGATTCTGTTGTGAATTGTAAACAAAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAAGTGC
 AATTTAGAAGTACATGGGAGGTGATTAGCAATTGTGAGGATTTAAAAACACCATAACCCATGGTGACACCACCT
 CCTCCACCTGTCTCTCATTGCTGAAGATCAGTCAGAACAGAATTGTGTGCTTAGTTCTGATAAGTCTGGAAGCATG
 GGGGGTAAGGACCCTAAATCGAATGAATCAAGCAGCAAACATTCTCTGCTGCAGACTGTTGAAAATGGATCC
 TGGGTGGGATGGTCACTTGATAGTACTGCCACTATTGTAATAAGCTAATCAAATAAAAGCAGTGTGATGAA
 AGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACCTCCATCTGCTCTGGAATTAAATATGCA
 TTTCAGGTGATTGGAGAGCTACATCCCAACTCGATGGATCGAAGTACTGCTGACTGATGGGAGGGATAAC
 ACTGCAAGTTCTGTATTGATGAAGTGAACAAAGTGGGCCCCATTGTCATTGTTATGCTTGGGAAGAGCTGCT
 GATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTGTTATGTTAGTCAAGATGAAGCTCAGAACAT
 GGCTCATTGATGCTTGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAACAGTCTCCCTCAGCTCGAAAGT
 AAGGGATAACACTGAATAGTAATGCTGGATGAACGACACTGTCTGCTCCCAGTATTCTCTCTGGGATCCCAGTGGAAACAATAATGGAAAAT
 TTCAACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCCAGGAACGACTGCAAAGGTGGCACTTGGGCATAC
 AATCTTCAGCCAAGCGAACCCAGAAACATTAACATTACAGTAACCTCTCGAGCAGCAAATTCTTCTGTGCCT
 CCAATCACAGTGAATGCTAAAATGAATAAGGACGAAACAGTTCCCCAGCCAAATGATTGTTACGCAGAAATT
 CTACAAGGATATGTAACCTGTTCTGGAGCCAATGTGACTGCTTCATTGAATCACAGAACATAGAACAGTT
 TTGGAACTTTGGATAATGGTGCAGGCCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTACAGCA
 TATACAGAAAATGGCAGATATAGCTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGG
 CCTCCACTGAATAGAGCCCGTACATACCAGGCTGGTAGTGAACGGGAAATTGAAGCAAACCGCCAAGACCT
 GAAATTGATGAGGAACTCAGACCACCTGGAGGATTTCAGCCGAACAGCATCCGGAGGTGCATTGTTGGTATCA
 CAAGTCCCAAGCCTCCCTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCACAGTCATGAG
 GATAAGATTATTCTTACATGGACAGCACCAGGAGATAATTGATGTTGGAAAAGTTCAACGTTATATCATAAGA
 ATAAGTGCAGTATTCTGATCTAAGAGACAGTTGATGATGCTCTCAAGTAATAACTACTGATCTGTCACCA
 AAGGAGGCCAACTCCAAGGAAAGCTTGCAATTAAACAGAAAATCTCAGAAAGAAAATGCAACCCACATATT
 ATTGCCATTAAAAGTATAGATAAAAGCAATTGACATCAAAGTATCCAACATTGCAACAGTAACCTTGTTC
 CCTCAAGCAAATCTGATGACATTGATCCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAAT
 TCTGGAGTTAATATTCTACGCTGGTATTGTCGTGATTGGGTCTGTTGTAATTGTTAACTTTATTAAAGTAC
 ACCATT**TGA**ACCTTAACGAAGAAAAAAATCTCAAGTAGACCTAGAAGAGAATTAAAAACAAATGTA
 GTAAAGGATATTCTGAATCTAAAATCATCCCATGTGTGATCATAAAACTCATAAAAATAATTAAAGATGTC
 GAAAAGGATACCTGATTAAATAAAACACTCATGGATATGTAAGAAACTGTCAAGATTAAATTAATAGTTCA
 TTTATTGTTATTGTAAGAAATAGTGTGATGAAACAAAGATCCTTTCTACTGATACCTGGTTGTATATT
 ATTGATGCAACAGTTCTGAAATGATATTCAAAATTGATCAAGAAATTAAATCATCTATCTGAGTAGTC
 AATACAAGTAAAGGAGAGCAAATAAACACATTGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

71/168

FIGURE 71

CTCCTTAGGTGGAAACCCCTGGGAGTAGAGTAACGTACAGCAAAGACCGGAAAGACCATACTGCCCGGGCAGGG
 TGACAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCTCCTATTCAAGGAAAGACGCCAAGGTAATT
 GACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTCCCTCTGAACCCCCAGTTATGCCAGGATTACTAG
 AGAGTGTCAACTCAACCAGCAAGCGGCTCCTCGGCTTAACCTGTGGTGGAGGAGAGAACCTTGCGGGCTGC
 GTTCTCTTAGCAGTGTCAAGGTGACTTGCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCTCTGCTGTG
 GCTGCACATCAGGAAGGCTGTGATGGAATGAAGGTGAAAACCTGGAGATTCACCTCAGTCATTGCTCTGCCT
 GCAAGATCATCCTTAAAGTAGAGAAGCTGCTCTGTGTTAGCTCAAGAGGGCAGAACTCGTTCTAGAA
 GGAAATGGATGCAAGCAGCTCGGGGGCCCCAACGCATGCTCCTGTGGCTAGCCCAGGGAAAGCCCTCCGTG
 GGGGCCCCGGCTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCCTGA**ATG**ATGATGGTTGCCGGGG
 CTGCTTGCGTGGATTCCCGGGTGGTTCTGGTGTGCTCTCTGCTGTGCTATCTCTGTCCTGTACATGTTG
 GCCTGCACCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAAGGAGGGTAC
 CAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCGCAACTACGTGAGCAGCCTGAAGCGCAGATCGCACAGCTC
 AAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAATGGCAGTACCAAGCCAGCAGTGTGCTGGCTGGGT
 CTGGACAGGAGCCCCCAGAGAAAACCCAGGCCACCTCCTGGCCTTCCTGCACTCGCAGGTGGACAAGGGAG
 GTGAATGCTGGCGTCAAGCTGCCACAGAGTATGCAAGCAGTGTGCTTCTGATAGCTTACTCTACAGAAGGTGTAC
 CAGCTGGAGACTGGCTTACCCGCCACCCGGAGGAGAAGGCTGTGAGGAAGGACAACGGGGATGAGTTGGTGGAA
 GCCATTGAATCAGCCTTGGAGACCCCTGAACATCCTGCAAGAACAGCCCCAATCACCGCCTTACACGGCCTCT
 GATTTCATAGAAGGGATCTACCGAACAGAAAAGGGACAATTGTGATGAGCTCACCTTAAAGGGGACCAC
 AAACACGAATTCAAACGGCTCATCTTATTCGACCATTCAGCCCCATCATGAAAGTGA AAAAATGAAAAGCTAAC
 ATGGCCAACACGCTTATCAATGTTATCGCCTCTAGCAAAAGGGTGGACAAGTCCGGCAGTTATCGCAGAAT
 TTCAGGGAGATGTGCAATTGAGCAGGATGGGAGAGTCCATCTCACTGTTACTTTGGAAAGAAGAAATAAAT
 GAAGTCAAAGGAATACTTGA AAAAACACTTCAAAGCTGCCACTTCAGGAACTTAACCTTACCCAGTGAATGGA
 GAATTTCTGGGAAAGGGACTTGATGTTGGAGCCGTTCTGGAAGGGAAGCAACGTCTCTCTTTCTGT
 GATGTGGACATCTACTCACATCTGAATTCTCAATACGTGAGGCTGAATACACAGCCAGGGAAAGGATTT
 TATCCAGTTCTTCAGTCAGTACAATCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTGGAAACAG
 CAGCTGGCATAAAGAAGGAAACTGGATTGGAGAGACTTGGATTGGGATGACGTGTCACTATCGGTACAG
 TTCATCAATATAGTGGTTGATCTGGACATCAAAGGCTGGGGCGAGAGGATGTGCACTTCACTCTGGCATGAG
 GACGAGCTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTG
 GGCATGCTGGTGTCAAGGACAGAGATAGGGCTCACCTCGAACAGAAAACAGAAAGACAAGTAGCAAAAAAAC
TGAACTCCCAGAGAAGGATTGGGAGACACTTTCTTCCTTGCAATTACTGAAAGTGGCTGCAACAGAGA
 AAAGACTTCCATAAAGGACGACAAAAGAATTGGGACTGATGGGTAGAGATGAGAAAGCCTCGATTCTCTGT
 TGGGCTTTACAACAGAAATCAAATCTGGCTTGTGCAAAAGAACCCAGTTGCACTTGTGAAGTGTCT
 GACAAGGCAGAATGCTGTGAGATTATAAGCCTAATGGTGGAGGTTTGATGTTACAATACACTGAGA
 CCTGTTGTTGTGCTCATGAAATATTGATGATTAAAGCAGTTGTAAGGAAATTCTAGCATGAAAG
 CAAGCATATTCTCTCATATGAACTGAGCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAAATATCAGAAG
 CAGGAGAGGAGATAGGCTTATTATGATACTAGTGAAGTACATTAAGTAAAATGGACCAGAAAAGAAAAGA
 ACCATAAATATCGTGTATTTCCCCAAGATTAACCAAAAATCTGTTATCTTTGGTTGCTTTAA
 CTGCTCCGTTTTCTTATTTAAAGTGCACCTTTCTGAGTTAGTCTGCTTATTAAATTAC
 CACTTGCAGCCTACAAGAGAGCACAAGTGGCCTACATTATTTATTTAAAGAAGATACTTGAGATGCA
 TTATGAGAACTTCAGTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGCAGGC
 ACTGAATGTCAGGCATTGAGACATAGGGAGGAAGGATGGTTGACTAATACAGACGTACAGATACTTCTCTGAAG
 AGTATTTCAGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGTTACAGAAAAGGAAACT
 CATTCACTGGTGTATCGTGTACCTAAAAGTCAGAACCCACATTTCTCTCAGAAGTAGGGACCGCTTT
 CTTACCTGTTAAATAACCAAAGTATACCGTGTGAACCAAACATCTTTCAAAACAGGGTGTCTCCTCTGG
 CTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATTGTAAGGATCAATCCATCTG
 CCAGAACATCTAGTGGAGATGGAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAAAGTAAC
 TTAATTAAAGCAGTTCTACTCAATCAGAACAGATGCTTCTGAAAATTGCAATTACATTCAA
 TAAAATAAACAGTTAACATAGTGGTTCTGCTCAGTAACACTGTTAAAAGCTCAAGAACATTCAA
 CTAATTATCTCTTGAGTCCTGCTTCTGTTGCTCAGTAACACTGTTAAAGCTCAAGAACATTCAA
 GCTGTTGGTGTGTTAAAAAATGCATTGATTGATTTGACTGGTAGTTATGAAATTAAATTAAACACAGGCCA
 TGAATGGAAGGTGGTATTGCACAGCTAATAAAATGATTGTTGGAATATGAA

73/168

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCCTGGGGATC
CAGAAACCCATGATAACCTACTGAACACCGAATCCCCTGGAAGGCCACAGAGACAGAGACAGC
AAGAGAACAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCA
CTCCTCCCTCCCTCTCTCTGCCTGTCTAGTCCTCTAGTCCTCAAATTCCAGTCCCCTGC
ACCCCTTCCTGGGACACTATGTTGTTCTCCGCCCTCTGCTGGAGGTGATTGGATCCTGGCT
GCAGATGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCTCT
TACCCCTGAGTGTGGAAACAATGCCAGTCGCCATCGATATTGACACAGACAGTGTGACATT
GACCCTGATTGCCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGAC
CTGCACAACAATGCCACACAGTGCAACTCTCTGCCCTCACCTGTATCTGGTGGACTT
CCCCGAAAATATGTAGCTGCCAGCTCCACCTGCACTGGGTCAGAAAGGATCCCCAGGGGG
TCAGAACACCAAGATCAACAGTGAAGCCACATTGCAAGAGCTCCACATTGTACATTGACTCT
GATTCCATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCTGCCTGGCATC
CTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTGAGTCACGGCATGAA
GTCAGGCATAAGATCAGAACCTCAGTGCCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAA
CAGCTGGGCAGTACTCCGCTACAATGGCTCGCTCACAACTCCCCCTGCTACCAGAGTGTG
CTCTGGACAGTTTTATAGAAGGTCCCAGATTCAATGGAACAGCTGGAAAAGCTTCAGGG
ACATTGTTCTCACAGAACAGAGGAGCCCTCAAGCTCTGGTACAGAAACTACCGAGCCCTCAG
CCTCTCAATCAGCGCATGGCTTGCCTTCAATCCAAGCAGGATCCTCGTATACCACAGGT
GAAATGCTGAGTCTAGGTGTAGGAATCTGGTTGGCTGTCTGCCTCTGGCTGTTAT
TTCATTGCTAGAAAGATTGGAAGAACAGAGGCTGGAAAACGAAAGAGTGTGGCTTCACCTCA
GCACAAGCCACGACTGAGGCATAAATTCCCTCTCAGATAACCATGGATGTGGATGACTCCCTT
CATGCCTATCAGGAAGCCTCTAAATGGGTGTAGGATCTGCCAGAAACACTGTAGGAGTAG
TAAGCAGATGTCCCTCCCTGGACATCTCTTAGAGAGGAATGGACCCAGGCTGTCAATTCC
AGGAAGAACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGAATGAGGAAATCGCTGTG
TTGTTAATGCAGAGANAAACTCTGTTAGTGCAAGGGAAAGTTGGATATACCCCAAAGTC
CTCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCAGGATCTCTCCTAGGATAA
AGAGTTGCTGTTGAAGTTGTATATTTGATCAATATATTGAAATTAAAGTTCTGACTTT

75/168

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGCGCCTGGGGACGGGCAGTCCCTGTG
TCTCTGGTGGTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATG**TCC
TACAATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCA
TCACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCCTG
ACAGCTCCAGAGAAGTGGAAAGAGAAATCCAGAACCTCCTGTTCCATGCAACAAATATA
TCCAATCTGAAGTATAACGTGTCTGTGAATACTAAATCAAACAGAACGTGGTCCCAGTGT
GTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTG
GAGTCCTCGTCCCAGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACT
TTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTCTGGTATGTTGCCATATCT
ATTACCGTGTCTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAA
GAGAACACCCAGCAAATTGATTGATTGAAATGAATTGACAAAAGATTCTTGTG
CCTGCTGAAAAAATCGTGATTAACTTATCACCTCAATATCTGGATGATTCTAAAATTCT
CATCAGGATATGAGTTACTGGAAAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCC
AGCGGGAACCTGAGGCCCTCAGGAGGAAGAGGAGGTGAAACATTAGGTATGCTTCG
TTGATGGAAATTGGACTCTGAAGAAAACACGGAAGGTACTTCTCAGCAGCAAGAG
TCCCTCAGCAGAACAAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAAC
ACTGACATTGTGCGGGCCTGAAGAGCAGGAGCTCAGTTGCAGGAGGAGGTG
GGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTGGCCCGCAAACGTTACAGTACTCA
TACACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCACACAGACTCGGAGGG
CCGGAGGAAGGCCATCGACGACCCCTGGTCAGTGGATCCCCAAACTGGCAGGCTGTG
CCTTCGCTGTCCAGCTCGACCAGGATTCAAGAGGCTGCAGCCTCTGAGGGGATGGC
GGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGA
AATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATATGTGCAGATGGAAA**A**
TGCCAAACACTTCTTTGCCTTGTGAAACAAAGTGAGTCACCCCTTGATCCCA
GCCATAAAAGTACCTGGATGAAAGAAGTTTCCAGTTGTCAGTGTCTGTGAGAATTACTT
ATTTCTTCTATTCTCATAGCACGTGTGATTGGTTATGCATGTAGGTCTCTAACAA
TGATGGTGGGCTCTGGAGTCCAGGGCTGGCGGTTGTTCTATGCAGAGAAAGCAGTCAATA
AATGTTGCCAGACTGGGTGCAGAATTATTGAGTGGGTGT

77/168

FIGURE 77

GAGGAGCAGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCCTGCAC TTGCTGCCCTCTGAC
ACCTGGGAAGATGGCCGGCCC GTGGACCTTCACCCTTCTCTGTGGTTGCTGGCAGCCACCTT
GATCCAAGCCACCCCTCAGTCCC ACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAA
GCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAG
TGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCT
GAAGCACATCATCTGGCTGAAGGT CATCACAGCTAACATCCTCAGCTGCAGGTGAAGCCCTC
GGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCC
CCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGA
CACCAAGTGCAAGTGGCCCCACCCGCCTGGCCTCAGTGACTGTGCCACCAGCCATGGGAGCCT
GCGCATCCA ACTGCTGTATAAGCTCTCCTCCTGGTGAACGCC TAGCTAACAGCAGGT CATGAA
CCTCCTAGTGCCATCCCTGCCAATCTAGTGAAAACCAGCTGTGTCCC GTGATCGAGGCTTC
CTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGA
CCGTCTGGAGTTGACCTTCTGTATCCTGCCATCAAGGGT GACACCATT CAGCTCTACCTGGG
GGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCT
GACAATGCCCACCCCTGGACAACATCCC GTTCAGCCTCATCGTGAGTCAGGACGTGGT GAAAGC
TGCAGTGGCTGCTGTGCTCTCCAGAAGAATT CATGGCCTGTTGGACTCTGTGCTTCCCTGA
GAGTGCCCATCGGCTGAAGTCAAGC ATCGGGCTGATCAATGAAAAGGCTGCAGATAAGCTGGG
ATCTACCCAGATCGTGAAGATCTTA ACTCAGGACACTCCGAGTTTTATAGACCAAGGCCA
TGCCAAGGTGCCCAACTGATCGTGTGCTGGAAAGTGTTCCTCCAGTGAAGCCCTCCGCCCTT
GTTCACCCCTGGCATCGAAGCCAGCTCGGAAGCTCAGTTTACACCAAAGGTGACCAACTTAT
ACTCAACTTGAAATAACATCAGCTCTGATCGGATCCAGCTGATGAACTCTGGGATTGGCTGGTT
CCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCACTCCATCCTGCTGCCAACCA
GAATGGCAAATTAAAGATCTGGGTCCCAGTGT CATTGGTGAAGGCCTTGGGATT CGAGGCAGC
TGAGTCCTCACTGACCAAGGATGCCCTTGCTTACTCCAGCCTCCTGTGAAACCCAGCTC
TCCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAAGGCTGGTCCCAGCTGGAGT
ATGGGTGTGAGCTCTATAGACCATCCCTCTGCAATCAAAACACTTGCTGTGAAAAAA

79/168

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA
GCTTCAGCCTGAAGACAAAGGGAGCAGTCCTGAAGACGCTCTACTGAGAGGTCTGCC**ATGGC**
CTCTCTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGCTTGGGCACACTGGT
TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGT
TGGCTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA
CATCTATAGCACCCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGAC
ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTT
CTGCCAGGAATCCCGAGCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTTGG
AGGCCTCCTGGGATTCAATTCTGTTGCCTGGAATCTCATGGGATCCTACGGGACTTCTACTC
ACCACTGGTGCCTGACAGCATGAAATTGAGATTGGAGAGGGCTTTACTGGGCATTATTC
TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTCTGCTCATCCCAGAGAAATCG
CTCCAACTAACGATGCCTACCAAGCCAACCTCTGCCACAAGGAGCTCCAAGGCCTGG
TCAACCTCCAAAGTCAAGAGTGAGTTCAATTCTACAGCCTGACAGGGTATGTG**TGA**AGAAC
CAGGGGCCAGAGCTGGGGGTGGCTGGTCTGTGAAAAACAGTGGACAGCACCCGAGGGCCA
CAGGTGAGGGACACTACCACTGGATCGTGTAGAAGGTGCTGCTGAGGATAGACTGACTTGG
CCATTGGATTGAGCAAAGGCAGAAATGGGGCTAGTGTAAACAGCATGCAGGTTGAATTGCCAA
GGATGCTGCCATGCCAGCCTTCTGTTCTCACCTGCTGCTCCCTGCCCTAAGTCCCC
AACCTCAACTTGAAACCCATTCCCTTAAGCCAGGACTCAGAGGATCCCTTGCCCTGGT
TTACCTGGACTCCATCCCAAACCCACTAATCACATCCACTGACTGACCCTCTGTGATCAA
AGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTATTGCTGGGATGGGAAGGAGAACAGT
GGCTTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAAACTGATTGGCCC
TGGAACCTCCATCCCACTTTGTTATGACTCCACAGTGTCCAGACTAATTGTGATGAACTG
AAATAAAACCATCCTACGGTATCCAGGAAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA
GGCAGCCTGGACATTTAAAAAAATA

81/168

FIGURE 81

CCACACGGTCCGCGCTCTCCCTCTGCTGGACCTCCTCGTCTCCATCTCCCTCCTT
TCCCCGCGTTCTCTTCCACCTTCTCTTCCACCTTAGACCTCCCTCCTGCCCTCCTT
TCCTGCCACCCTGCTGCTCCTGGCCCTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGT
CTGTGGGTTGATCTGTGGCCCTGTGCCTCCGTGTCCCTTCGTCTCCCTCCGACTCC
GCTCCGGACCAGCGGCCTGACCCCTGGGGAAAGGATGGTCCCGAGGTGAGGGTCCTCCCTC
CTTGCTGGACTCGCGCTGCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACAT
GTTCTGCCTTCCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGCC
ACAAGGCCTGATGTACTGCCTGCCTGTACCTGCTCAGAGGGCGCCATGTGAGTTACCG
CCTCCACTGTCCGCCTGTCCACTGCCCTGAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAA
GTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACACGG
GACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCTGCCCAA
CCAGTGTGTCCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCCCTCACACCTGCCCGA
ACCAGGCTGCCAGCACCCCTCCACTGCCAGACTCCTGCTGCCAGCCTGCAAAGATGAGGC
AAGTGAGCAATGGATGAAGAGGACAGTGTGCACTGCTCCATGGGTGAGACATCCTCAGGA
TCCATGTTCCAGTGTGCTGGAGAAAGAGAGGGCCGGCACCCCAGCCCCACTGGCCTCAG
CGCCCTCTGAGCTTCACTCCCTGCCACTTCAGACCCAAGGGAGCAGGCAGCACACTGTCAA
GATCGTCCTGAAGGAGAACATAAGAAAGCCTGTGCACTGGGGAAAGACGTACTCCCACGG
GGAGGTGTGGCACCCGGCTTCCGTGCCTCGGCCCTGCCATCCTATGCACCTGTGA
GGATGGCGCCAGGACTGCCAGCGTGTGACCTGCTCCACCGAGTACCCCTGCCGTACCCCGA
GAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAGGACAAAGCAGACCCCTGCCACAGTGA
GATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGTCTCGTCCACACATCGGTATCCCC
AAGCCCAGACAAACCTGCGTCGCTTGGCCTGGAACACGAGGCCTGGACTTGGTGGAGATCTA
CCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAGGTGAAGTACCTGGCCAAG
GCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAGGAAGCAAGACTCCAGA
AAGAGGCACAGCACTCCGACTGCTCGTGGCCCCACGAAGGTCAGTGGACGTCTCCTAG
CCCAGACCCCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGAC
CTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATTAATAATAAGAAGTTGC
ATTACCTCAAAAAAAAAAAAAAA

FIGURE 83

85/168

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCTAACCCCAGAC**ATG**CTGCTGCTGCTGCC
CTGCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCC
GTGACGGTGAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTGCATGGC
TGGATTACCTGGCCCAGTAGTCATGGCTACTGGTCCGGGAAGGGGCAATAACAGACCAG
GATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGAGGAGACTCGGGACCGATT
CACCTCCTGGGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGT
GATGCGGGGAGATACTTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATCAC
CGGCTCTCTGTGAATGTGACAGCCTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTG
GAGTCGGCTGCCCTCAGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCC
CCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCG
GTGCTCACCCCTCATCCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTCAAGGTGACCTTC
CCTGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCCTACCCGCTCAGAAC
TTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTGGAAATGGCTCATCT
CTGTCACTCCCAGAGGGCAGTCCTGCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAAT
CCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCCTGACCCCTGTGCCCTCACAGCCCTAAAC
CCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGAATTCACCTGCAGAGCT
CAGAACCCCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGCAGAGCAAAGCCACATCA
GGAGTGACTCAGGGGTGGTCGGGGAGCTGGAGGCCACAGCCCTGGCTTCCCTGTCCCTG
GTCATCTCGTTGAGGTCTGCAGGAAGAAATCGGCAAGGCCAGCAGCGGGCGTGGGA
GATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTCAGCCTCTCAGGGGCCCTGACTGAA
CCTTGGGCAGAACAGTCAGGCTCAGCAGGCTCAGCTTCCAGATGGTGAAGCCTTGGACTCGCGGGGA
GAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTGAAGCCTTGGACTCGCGGGGA
CAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAGA**TGA**GAAACTGCAGAGACT
CACCCCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCAGAGGCTGATTCTGTAG
AATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATGTGCAGAGTGAAGAAC
ACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCCTCCCTTTATT
TTTAACAAAAGACAGACAAATTCTCA

87/168

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGT
ACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAAACTCAGCTCCTGCTG
TTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGG
ACCTGTTCTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGT
GCATTTGATGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTCTGTGAC
ATGACCTCTGGGGTGGCGCTGGACCCCTGGTGGCCAGCGTCATGAGAATGACATGCGTGGG
AAGTGCACGGTGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGG
GACGGCAACTGGCCAACATACAACACCTTGGATCTGCAGAGGCGGCCACGAGCGATGACTAC
AAGAACCCCTGGCTACTACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAG
TCCCCCATGCAGCACTGGAGAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTC
CAGACACTGGACATAATCTGTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGA
AAGTGTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTGGCAGGCCAGAAA
ACAGCATCTTAACTCACCTATGCCAGCGGGATTCACTGCAGGGATTGTTCAAGTCAGGG
GTATTTAATAACGAGAGAGCAGCCAACGCCCTGTGTGCTGAATGAGGGTACCGGATGTAAC
ACTGAGCATCACTGCATTGGTGGAGGAGGATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGA
GATTTTCTGGTTTGATTGGAGTGGATATGGAACCTCATGTTGGTACAGCAGCAGCCGTGAG
ATAACTGAGGCAGCTGTGCTCTATTCTATCGT**TGA**GAGTTTGTGGAGGGAACCCAGACCT
CTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCAGTAGCTAGAATGTTAATG
GCAGAAGAGAAAACAATAATCATATTGACTCAAGAAAAAA

89/168

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAG
ACCGCCGCCCTTGTCCCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCT
GACGCTCCTGGCGCATCTGGTGGTCGTATCACCTTATTCTGGTCCCGGGACAGAACATACA
GGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCA
CGCGCTCTCTGTCAACCCTGGGCCTTTGCAGTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTC
CATGTTAACAGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCT
GTCCTTCTTCATATTGAGCGTTGGAGTGCACTA
CTGATTGGTACATTGTCTGCAG
TGCCCTTCCAGCTGTCACTGAAATGGCTTATTGTCACCGTCTTGGCTGAAAAAGAAACC
CTTC**TGA**TTACCTTCATGACGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGT
ATTCTGGAAGAAGGAAGGCATAGGCTTCGGTTTCCCCTCGAAACTGCTTCTGCTGGAGGA
TATGTGTTGGAATAATTACGTCTGAGTCTGGATTATCCGCATTGTATTTAGTGCTTGTAA
TAAAATATGTTTAGTAACATTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAA
AAA

91/168

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGACGAGAGCGGAGGAAGATGCAACTG
ACTCGCTGCTGCTTCGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGAT
GATGGTCCTCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCAGCCCCGGCCCCGG
GTGCCTCGGAAGCGGGGCCACATCTCACCTAACGCCCAGCCCCATGCCAATTCCACTCTCCTA
GGGCTGCTGGCCCCGCTGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCCAC
CACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTGCTGGGCAGTCTACTCCAAC
ATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGAAAGATTGTGGACCATGGCAATGGACC
TTCAGCGTCCACTTCAAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGC
CCCAGTAAAGCTGTAGAGTCCACCAGGAACAGCAGATCTCATCGAAGCCAAGGCCTCCAAA
ATCTTCAACTGCCGGATGGAGTGGAGAAGGTAGAACGGGCCGGACCTCGCTTGACC
CACGACCCAGCCAAGATCTGCTCCGAGACCGCTCAGAGCTCAGCCACCTGGAGCTGCTCC
CAGCCCTCAAAGTCGTCTGTCTACATCGCCTCTACAGCACGGACTATCGCTGGTCCAG
AAGGTGTGCCAGATTACAACCTACCATAGTGTACGACCCCTACTACCCATCTGGGTGACCC
GCAGGCCACAGAGGCCAGGCCAGGGCTGGAGGACAGGCCCTGCCATGCAGGAGACCATCTGG
ACACCGGGCAGGGAAAGGGTTGGGCCTCAGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAA
GTGGGCCAGGCCAAGTCTCAAGTGGCAGAGAAAGGTCCAAGTGCTGGTCCAAACCTGAA
GCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGTGGCTCTGTGCAGCCTCACA
GGGCTTGCCACGGAGCCACAGAGAGATGCTGGTCCCCGAGGCCCTGGCAGGCCGATCAG
TGTGGCCCAGATCAAGTCATGGAGGAAGCTAACGCCCTGGTCTTGCCATCCTGAGGAAAG
ATAGCAACAGGGAGGGGAGATTTCATCAGTGTGGACAGCCTGTCAACTTAGGATGGATGGCT
GAGAGGGCTCTAGGAGCCAGTCAGCAGGGTGGGTGGGCCAGAGGAGCTCTCAGCCCTG
CCTAGTGGCGCCCTGAGCCCTGTGCTGCTGAGCATGGCATGAGGCTGAAGTGGCAACC
CTGGGTCTTGATGTCTTGACAGATTGACCATGTCTCCAGCCAGGCCACCCCTTCCAAA
ATTCCCTCTGCCAGTACTCCCCGTACCAACCCATTGCTGATGGCACACCCATCCTTAAG
CTAACAGGACGATTGTGGCTCCCACACTAACGCCACAGCCATCCGCGTGTGTGTC
CCTCTCCACCCCAACCCCTGCTGGCTCTGGAGCATCCATGTCCGGAGAGGGTCCCT
CAACAGTCAGCCTCACCTGTCAGACCGGGTTCTCCGGATCTGGATGGCGCCGCCCTCAG
CAGCGGGCACGGTGGGGGGCCGGCGAGAGCATGTGCTGGATCTGTGTGTCT
GTCTGTGGGTGGGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTTGTGTGA
AGAATCGTGTCTTGAGCAGGAAATAAGCTTGCCCCGGGGCA

93/168

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTCGGCTGCCCTTCATTGCCTCGGGCCTGCGCTCG
CCCTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCATGCCGGAGCTT
TCTTCTGGTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTG
ACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTCTATA
TCCAAGAAATGTTCCGATTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGA
GTATAAACCCAGGTGAGACAGCACCCCTATGCGACTGCTGCCATGTTCTGGCTGGGCT
TTGGAATCATGAGTGGAGTATTCCTTGTGAATACCCTATCTGACTCCTGGGCCAGGCA
CAGTGGCATTGAGATTCTCCTCAATTCTCCTTATTGAGCTTCATGACGCTGGTCA
TTATCTGCTGCATGTATTCTGGGCATTGTATTTGATGGCTGTGAGAAGAAAAAGTGGG
GCATCCTCCTATCGTCTCCTGACCCACCTGCTGGTGTAGCCCAGACCTCATAAGTTCTT
ATTATGGAATAAACCTGGCGTCAGCATTATAATCCTGGTGTGCTGGCACCTGGCATTCT
TAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTTC
TTCTTTACAACCAGCGCTCCAGATAACCTCAGGGAACCAGCACTCCAAACCGCAGACTACA
TCTTAGAGGAAGCACAACGTGCCTTTCTGAAAATCCCTTTCTGGTGGATTGAGAAA
GAAATAAAACTATGCAGATA

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCCTCG
TGGACCCAAAGGTAGCAATCTGAAACATGAGGGAGTACGATTCTACTGTTTGTCTTAGGAT
CAAACTCGGTCATTACCAACAGCTCAAACCTGCTTGGACTCCCTCCCACAAAACGGCTCCGG
ATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATAC
CATTAACACAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGA
CACCTGGTACCCAGACCCACCCATTGACCCCTGGAGGGTTGAATGTACAACAGCAACTGCACC
CACATGTGTTACCAATTTCACGAGCCTCATCCATTGTTCCGGAGGCATCCTGCCA
CCAGTCAGGCAGGGCTAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTG
TAAATCCTGCCACCCAGGGAACCCAGCAGGCCGCCTCCAACTCCAGTGGCACAGATGACG
ACTTTGCAGTGACCACCCCTGCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGGCCACCA
CAGAATCAGCAAATGGAATTCAGTAAGCTGTTCAAATTTCAAACTAAGCTGCCTCGAATT
TGGTGATACATGTGAATCTTATCATTGATTATATTGGAATAGATTGAGACACATTGGATA
GTCTTAGAAGAAATTAAATTCTTAATTACCTGAAAATATTCTTGAAATTTCAGAAAATATGTT
CTATGTAGAGAATCCAACTTTAAAAACAATAATTCAATGGATAATCTGTCTTGAAATAT
AACATTATGCTGCCCTGGATGATATGCATATTAAAACATATTGGAAAATGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

97/168

FIGURE 97

GCTCAAGTGCCCTGCCTGCCCAACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCT
CTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGCGCTGGAGGGCTGTCTGAC
CATGTCCCTGCCTGGCTGTGGCTGCTTGTCTCCGTCCCCCAGGCTCTCCCAAGGCCA
GCCTGCAGAGCTGTCTGGAAAGTCCAGAAAATGCTGGAAATTCCCTTATACCTGAC
CAAGTTGCCGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTCTGCTGGACTCAGGCAA
GGCAACTGAGGGCCATTGCTATGGATCCAGATTCTGGCTCTGCTGGTACCCAGGCT
GGACCGAGAGGAGCAGGCAGAGTACCAAGTCACAGGTACCCCTGGAGATGCAGGATGGACATGT
CTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCATTT
CTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGTACCAAGGCCTGGCATCCCTCCT
CCTTGAGGCTTCAGACCAGGATGAGCCAGGCACAGCCAACTCGGATCTCGATTCCACATCCT
GAGCCAGGCTCCAGCCCAGCCTCCCCAGACATGTTCCAGCTGGAGGCTCGGCTGGGGCTCT
GGCCCTCAGCCCCAAGGGGAGCACCCAGCCTGACCACGCCCTGGAGAGGACCTACAGCTGT
GGTACAGGTCAAGGACATGGGTGACCAGGCCAGGCCACTGCCACCGTGGAGT
CTCCATCATAGAGAGCACCTGGGTGCTCCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGT
CCTATAACCGCACCACATGGCCCAGGTACACTGGAGTGGGGTGATGTGCACATCACCTGGA
GAGCCATCCCCCGGGACCCTTGAAGTGAATGCAGAGGAAACCTCACGTGACCAGAGAGCT
GGACAGAGAACGCCAGGCTGAGTACCTGCTCCAGGTGCAGGCTCAGAACATGGCAGGAA
CTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAATGACAACGTGCCTATCTG
CCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCACCAGGTACTGAAGTGA
ACTGTCAAGCAGAGGATGCAGATGCCCGGCTCCCCAATTCCACGTTGTATCAGCTCCT
GAGCCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCTCCAGGTGGACCCACTCAGGCAG
TGTGACGCTGGGGGTGCTCCACTCCGAGCAGGCCAGAACATCCTGCTCTGGTCTGGCAT
GGACCTGGCAGGCCAGAGGGTGGCTTCAGCAGCACGTGTGAAGTCGAAGTGCAGTCACAGA
TATCAATGATCACGCCCTGAGTTCATCACTCCAGATTGGCCTATAAGCCTCCCTGAGGA
TGTGGAGCCGGGACTCTGGTGGCATGCTAACAGCCATTGATGCTGACCTCGAGCCGCCT
CCGCCTCATGGATTTCGCCATTGAGAGGGGAGACACAGAACGGACTTTGGCCTGGATTGGGA
GCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAGGCAGCTCCAAG
TCATGAGGTGGTGGTGGTGCAGAGTGTGGCAAGCTGGTGGGGCAGGCCAGGCCCTGG
AGCCACCGCCACGGTACTGTCCCCATCAGTCCCCAGCCGGCTTTCTGCTGACCATCCAGCC
CTCCGACCCCATCAGCCGAACCTCAGGTCTCCCTAGTCATGACTCAGAGGGCTGGCTCTG
CATTGAGAAATTCTCCGGGGAGGTGCACCCGCCAGTCCTGACTCTGCCCCCTGTGCCCTCCA
CACCTACACGGTCTGTGGAGGCCAGGATACAGCCTGACTCTGCCCCCTGTGCCCTCCA
ATACCTCTGCACACCCGCCAAGACCATGGCTGATCGTGAGTGGACCCAGCAAGGACCCGA
TCTGGCCAGTGGCACGGCTACAGCTCACCCCTGGTCCAACCCACGGTGCAACGGGA
TTGGCGCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGTGG
GCCACGTGAACACATAATCCCCTGGTGGCAGCCACAATGCCAGATGTGGCAGCTCTGGT
TCGAGTGATCGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGGGCATGAA
GGGCATGCCACGAAGCTGCGCAGTGGCATCCTGTAGGCACCCCTGGTAGCAATAGGAAT
CTTCCTCATCCTCATTTCAACCACTGGACCATGTCAAGGAAGAAGGCCGGATCAACCAGC
AGACAGCGTGCCCTGAAGGCAGTGT**TGA**ATGGCCCAGGCAGCTAGCTGGAGCTTGG
CTCTGGCTCCATCTGAGTCCCTGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGA
CAGAGTAGAAGCCCCCTCCATCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCT
GCAGAGCCTGGACACCAACTTTATGGACTGCCCATGGAGTGCCTCAAATGTCAGGGTGT
CCCAATAATAAGCCCCAGAGAACTGGCTGGCCCTATGGAAAAA
AAAAAAAAAAAG

99/168

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCC
TGAGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCC
TTTATCTCTCACCTTCAGTCCCTTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCT
TGGTACATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAG **TGA**AGATGCAGAA
AGGAAATGTTCTCCTTATGTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCCAATGA
GAATAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACGCCACCAA
CTCTGGGTCCAGTGTGACCTCCAGTGGGTCCAGCACAGCCACCATCTCAGGGTCCAGCGTGAC
CTCCAATGGGTCCAGCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGATCAGCAC
AGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTC
CAGCACACACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGG
GGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAA
CTCTGAGTCCACACAGTGTCCAGTAGGGCCAGCAGTGCACCAACTCTGAGTCTAGCACACT
CTCCAGTGGGCCAGCACAGCCACCAACTCTGACTCCAGCACACACTCCAGTGGGCTAGCAC
AGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTC
CAGCACAGTGTCCAGTAGGGCCAGCAGTGCACCAACTCTGAGTCCAGCACACACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCCAATGGGCTGGCACAGCCACCAA
CTCTGAGTCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAGT
GTCCAGTGGGCCAGCAGTGCACCAACTCTGAGTCCAGCACACACTCCAGTGGGCTAGCAC
AGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGACTC
CAGCACACACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGCCAACACAGCCACCAA
CTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCACCAACTCTGAGTCCAGCACAGT
GTCCAGTGGGCCAGCAGTGCACCAACTCTGAGTCCAGCACACACTCCAGTGGGCTAGCAC
AGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGACTC
CAGCACACACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGCCAACACAGCCACCAA
CTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGAATGCACACAA
TTCCCATAGTCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGGCCCTGGTGCCTGGGA
AATCTTCCCTCATCACCCCTGGTCTGGTTGTGGCGGCCGTGGGCTCTTGCTGGCTCTTCTT
CTGTGTGAGAAACAGCCTGTCCTGAGAAACACCTTAACACAGCTGTCTACCACCCCTCATGG
CCTCAACCCTGGCCTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCCAGGCCAGGTG
GAGTCCTAACTGGTTCTGGAGGAGACCAGTATCATGATAGCCATGGAGATGAGCGGGAGGAA
CAGCGGGCCCT**TGA**GCAGCCCCGGAAAGCAAGTGCCGATTCTCAGGAAGGAAGAGACCTGGGC
ACCCAAAGACCTGGTTCTTCTTCAATTCTCATCCCAGGGAGACCCCTCCAGCTTGTGAGATCCT
GAAAATCTTGAAGAAGGTATTCTCACCTTCTTGCCTTACCAAGACACTGGAAAGAGAATAC
TATATTGCTCATTAGCTAAGAAATAATACATCTCATCTAACACACAGACAAGAGAAGCT
GTGCTTGGCCCCGGGTGGGTATCTAGCTGAGATGAACTCAGTTAGGAGAAAACCTCCAT
GCTGGACTCCATCTGGCATTCAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

101/168

FIGURE 101

GGCGGGACGCCTCCCGTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTGTGACCCC
TACGGAGCCCCAGCTTGCACGCACCCCCACTCGCGTCGCCGCGCGTGCCTGCTTGTCA
GGTGGGAGGGCTGGAACATACAGGCTGAAAAACAGAGTGGGTACTCTCTCTGGGAAGCTGGCA
ACAAATGGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTTCTGAACCCAT
GGTCAATTAAACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT
TTGGAATCATGGTGTATGGAAAGGGATTACTTACTGACTCTGTTGGGAAGCTTT
TTGGAAGCATTTCATGCTGAGTCCCTTTACCTTGATGTTGTAAACCCATCTGGTATC
GCTGGATCAACAACCGCCTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCA
TGTTGGTGTAAAAGTGAATTATAACTGGGGATGCATTGTTCTGGAGAAAGAAGTGTCA
TCATGAACCACCGACAAGAATGGACTGGATGTTCTGTGAAATTGCCTGATGCGATATAGCT
ACCTCAGATTGGAGAAAATTGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGG
CCATGCAGGCTGCCTATATCTCATTCAAGGAAATGGAAGGATGACAAGAGCCATTG
AAGACATGATTGATTACTTTGTGATATTCAACGAACTTCAACTCCTCATATTCCAGAAC
GGACTGATCTCACAGAAAACAGCAAGTCTGAAGTAATGCATTGCTGAAAAAAATGGACTTC
AGAAATATGAATATGTTTACATCCAAGAAACTACAGGCTTACTTTGTGGTAGACCGTCTAA
GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACAAACATC
AATCAGAGAACGACCTCCTCCAAGGAGACTTCCCAGGGAAATCCACTTCACGTCCACCG
ATCCAATAGACACCCCTCCCCACATCCAAGGAGGACCTCAACTCTGGGCCACAAACGGTGG
AAGAGAAAAGAGAGAGGCTGCCTCTATCAAGGGAGAAGAATTTTATTTACCGGAC
AGAGTGTCAATTCCACCTTGCAAGTCTGAACTCAGGGCCTTGTGGTCAAATTGCTCTATAC
TGTATTGGACCCCTGTTGCCCTGCAATGTGCCTACTCATATATTGTACAGTCTTGTAA
GGTATTTATAATCACCATTGTAATCTTGTGCTGCAAGAGAGAATATTGGTGGACTGGAGA
TCATAGAACTTGCATGTTACCGACTTTACACAAACGCCACATTAAATTCAAAGAAAAATG
AGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTGGAAATGTTCTAAACCTT
TCTAAGCTCAGATGCATTGGCATGACTATGTCGAATATTCTTACTGCCATCATTATTGT
TAAAGATATTGCACTTAATTGTGGAAAATATTGCTACAATTTTTAATCTCTGAA
TGTAATTGCAACTGTGTACATAGCAGGGAGTGATGGGGTGAATAACTGGGCCAGAATA
TTATTAAACAATCATCAGGCTTTAAA

103/168

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCACGGCTCTGCAGACAGCTGGCCTGA
CCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTCATAGTGTGAGATCAACCCACAGG
AATATCC**ATG**GCTTTGTGCTCATTGGTCTCAGTTCTACGAGCTGGTGTCAAGGACAGTG
GCAAGTCACTGGACCAGGGCAAGTTGTCAGGCCTGGTGGGGAGGACGCCGTGTTCTCCTG
CTCCCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGTTCTCAGGAATCAGTTCCA
TGCTGTGGTCCACCTCTACAGAGATGGGAAGACTGGGAATCTAACAGATGCCACAGTATCG
AGGGAGAACTGAGTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTCTAACAGGCTAAAAAA
CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTCAGTTCCAGATTACGATGAGGA
GGCCACCTGGGAGCTGCAGGCTGGCAGCACTGGCTCACTTCTCTCATTTCCATCGTGGGATA
TGTTGACGGAGGTATCCAGTTACTCTGCCTGCTCAGGCTGGTCCCCCAGCCCACAGCCAA
GTGGAAAGGTCCACAAGGACAGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGTACAG
CCTGTATGATGTGGAGATCTCATTATAGTCCAGGAAAATGCTGGGAGCATAATTGTGTTCCAT
CCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTCCA
GCCCTCACCTTGGCGCCTGGCTCTATTACTCGGGTTACTCTGTGGTGCCTGTGGTGT
TGTATGGGATGATAATTGTTCTTCAAATCCAAAGGGAAAATCCAGGCGGAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCT
GGATCCAGAGACGGCTACCCGAAGCTCTGCCTTGATCTGAAACTGTAACCCATAGAAA
AGCTCCCCCAGGGAGGTGCCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCA
GGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGACAAAATGTAGGGTGGTATGT
GGGAGTGTGTGGGATGACGTAGACAGGGGAAGAACATGTGACTTGTCTCCAAACAATGG
GTATTGGGTCTCAGACTGACAACAGAACATTGTATTCACATTCAATCCCCATTATCAG
CCTCCCCCCCAGCACCCCTCTACACGAGTAGGGGTCTCCTGGACTATGAGGGTGGACCAT
CTCCTTCTCAATAACAAATGACCAGTCCCTTATTATAACCTGCTGACATGTCAGTTGAAGG
CTTGTGAGACCCATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCATATTCA
ATGTCCAGTGTCTGGGA**TGA**GACAGAGAAGACCTGCTAAAGGGCCCACACCACAGACC
CAGACACGCCAAGGGAGAGTGCCTCCGACAGGTGGCCCCAGCTCCTCTCCGGAGCCTGCGC
ACAGAGAGTCACGCCCAACTCTCCTTAGGGAGCTGAGGTTCTGCCCCGAGCCTGCA
GCAGCAGCAGTCACAGCTCCAGATGAGGGGGATTGGCCTGACCCCTGTGGAGTCAGAACCC
ATGGCTGCCCTGAAGTGGGAGGAATAGACTCACATTAGGTTAGTTGTGAAAACCTCCATC
CAGCTAACGCGATCTTGAACAAGTCACAACCTCCAGGCTCCTCATTGCTAGTCACGGACAGT
GATTCTGCCTCACAGGTGAAGATTAAGAGACAAACGAATGTGAATCATGCTTGAGGTTGA
GGGCACAGTGTGCTAATGATGTGTTTATATTACATTTCACCCATAAACTCTGTTT
GCTTATTCCACATTAATTACTTTCTCTACCAAATCACCATGGAATAGTTATTGAACAC
CTGCTTGTGAGGCTCAAAGAATAAGAGGAGGTAGGATTTCACTGATTCTATAAGCCCAG
CATTACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCA
TATCCCTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAT
ATATTAAAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAA
TATTTAAATATCAACCAGTGTAAATTGAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAA
AAAAAAAAAA

105/168

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTAGGGCTAG
GAAAAGAGTTGGAAACCCTGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCCT
GGCAGTGTGCATTGGACTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATT
CTATAGCACATTGTCATTACAAC TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAA
CAATTTACAGAAATGAGCCAGAGACTTGAATCAATGGTAAAAATGCATTTATAAACTCC
ATTAAGGGAAGAATTGTCAAGTCTCAGGTTATCAAGTTCACTAACAGAAGCATGGAGTGT
GGCTCATATGCTGTTGATTGTAGATTCACTCTACTGAGGATCCTGAAACTGTAGATAAAAT
TGTTCAACTGTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAGATCCTCA
CTCAGTTAAAATAAAAAAATCAACAAGACAGAACAGACAGCTATCTAAACCATTGCTGCGG
AACACGAAGAAGTAAAACCTAGGTCAAGGTCTCAGGATCGTTGGTGGACAGAAGTAGAAGA
GGGTGAATGCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGTGGAGAACCTT
ATTAATGCCACATGGCTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCTGCCAG
ATGGACTGCTCCCTTGGAGTAACAATAAAACCTCGAAAATGAAACGGGTCTCCGGAGAAT
AATTGTCCATGAAAATACAACACCCATCACATGACTATGATATTCTCTTGAGAGCTTTC
TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTCTCCCTGATGCATCCTATGAGTT
TCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCAGTGGAAATGATGGTTACAGTCA
AAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC TTGCAATGAACCTCAAGC
TTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGC
ATGCCAGGGTGA CTCGGAGGACACTGGTTAGTTCACTGCTAGAGATATCTGGTACCTTGC
TGGAAATAGT GAGCTGGGAGATGAATGTGCGAAACCCAAACAAGCCTGGTGGTTACTAGAGT
TACGGCCTGCGGGACTGGATTACTTCAAAAACTGGTATCT**AA**GAGACAAAAGCCTCATGGAA
CAGATAACATTTTTGTTTTGGGTGTGGAGGCCATTTAGAGATA CAGAATTGGAGA
AGACTTGCAAAACAGCTAGATTGACTGATCTCAATAAAACTGTTGCTGATGCATGTATTT
CTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTCTGCCAGATCAACTCTGTCATCTGTGA
GCAATAGTTGAAACTTATGTACATAGAGAAATAGATAATAACATTACATTACAGCCTGTA
TTCATTGTTCTCTAGAAGTTGTGAGAATTGACTGTTGACATAAAATTGTAATGCATA
TATACAATTGAAAGCACTCCTTCTTCAGTTCCCTCAGCTCCTCTCATTCAGCAAATATCCA
TTTCAAGGTGAGAACAGGAGTGAAAGAAAATAAGAAGAAAAATCCCTACATTTA
TTGGCACAGAAAAGTATTAGGTGTTCTTAGTGGAAATTAGAAATGATCATATTCTTAT
GAAAGGTCAAGCAAAGACAGCAGAACATACCAATCACTCATCTTAGGAAGTATGGGAAC TAA
GTTAAGGAAGTCCAGAAAGAACAGATATACTCTTATTTCACTTCAAACAAACTACTATG
ATAAAATGTGAAGAAGATTCTGTTTTGTGACCTATAATAATTACAAACTTCATGCAATG
TACTTGTCTAAGCAAATTAAAGCAAATATTATTTAACATTGTTACTGAGGATGTCAACATA
TAACAAATAAAATATAACCCCA

107/168

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCC
TGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGCGTGGGGCACCGGGCC
CAGCGCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTGGTCAAAGGATGGGGCTTC
TCCCTTACGGGGCTCACAAATGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTAC
GCCCTCAATCTGCTCTTGGTTAATGTCCATCAGTGTGGCAGTTCTGCTGGATGAGG
GACTACCTAAATAATGTTCTCACTTAACAGCAGAAACGAGGGTAGAGGAAGCAGTCATTTG
ACTTACTTCCCTGTGGTCATCCGGTCACTGATTGCTGTTGCTGTTCTTATCATTGTGGGG
ATGTTAGGATATTGTGGAACGGTAAAAGAAATCTGTTGCTTCTGCATGGTACTTGGAAGT
TTGCTTGTCAATTCTGTAGAACCTGGCTGTGGCGTTGGACATATGAACAGGAACATTATG
GTTCCAGTACAATGGTCAGATATGGTCACCTTGAAAGCCAGGGATGACAAATTATGGATTACCT
AGATATCGGTGGCTTACTCATGCTTGGAAATTTCAGAGAGAGTTAAGTGTGTGGAGTA
GTATATTCACTGACTGGTTGGAAATGACAGAGATGGACTGCCCTCAGATTCTGCTGTGTT
AGAGAATTCCCAGGATGTTCAAACAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAG
GGTTGTGGAAGAAAATGTATTCCCTTTGAGAGGAACCAAAACTGCAGGTGCTGAGGTTT
CTGGGAATCTCCATTGGGGTACACAAATCCTGGCCATGATTCTCACCATTACTCTGCTCTGG
GCTCTGTATTATGATAGAAGGGAGCCTGGACAGACCAATGATGTCCTGAAGAATGACAAC
TCTCAGCACCTGTCATGCCCTCAGTAGAACCTGGTAAACCAAGCCTGTCAAGAATCTTGAA
CACACATCCATGGCAAACAGCTTAAATACACACTTGAGATGGAGGAGTTTAAAAGAAATG
TCACAGAAGAAAACCACAAACTGTTTATTGGACTTGTGAATTGAGTACATACTATGTG
TTTCAGAAATATGAGAAATAAAATGTTGCCATAAAATAACACCTAACGATATACTATTCTA
TGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTCACCACCTGGACAATAATTGATGC
CCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCCTGTGTACTTTACTGAAC
ACAGTTATGTTGAGGCAGCATGGTTGATTGACATTCCGCATCCATGCAAACGAGTCACA
TATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAACTAGTATATAAGTACT
AATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATTACTCAGCGATCTATT
TTCTGATGCTAAATAAAATTATATCAGAAAACCTTCAATATTGGTACTACCTAAATGTGAT
TTTGCTGGTACTAAATATTCTACCACTTAAAGAGCAAGCTAACACATTGTCTTAAGCT
GATCAGGGATTTTGATATAAGTGTGTTAAATCTGTATAATTCACTGATTTCACTGTTCT
GATAATGTTAAGAATAACCATTATGAAAAGGAAAATTGCTGTATAGCATATTATTTA
GCCTTCCGTAAATAAGCTTACTATTCTGCTGGCTTATATTACACATATAACTGTTA
TTTAAATACTAACCAACTATTGAAAATTACCACTGTGATACATAGGAATCATTATTCA
ATGTAGTCTGGTCTTACCGAGTATTAATAAGAAAATTGACACATAACTTAGTTGATTGAGAA
AGGACTTGTATGCTGTTCTCCAAATGAAGACTCTTTGACACTAACACACTTTAAAA
AGCTTATCTTGCCTTCTCCAAACAGAACAGCAATAGTCTCCAAGTCAATATAATTCTACAGA
AAATAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAG
AGATTCTTGTAAATTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATT
TTTACAAGAGTATAGTATATTGAAATGGAAAAGTGCATTTACTGTATTTGTGTAT
TTGTTATTCTCAGAATATGGAAAAGAAAATTAAATGTGTCAATAAAATATTCTAGAGAG
TAA

109/168

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTATCCCACCTCATCCTCATCCTCTTCCTCTGATAAAGCCC
CTACCAGTGCTGATAAAAGTCTTCTCGTAGAGGCCTAGAGGCCTAAAAAAAAGTGCTTGA
AAGAGAAGGGGACAAAGAACCCAGTATTAAAGAGGATTTCCAGTGTCTGGCAGTGTCACC
CAGAAGG**ATG**CCTCCATTCTCGTCTCACCTGCCTCTCATCACAGGCACCTCCGTGTCACC
CGTGGCCCTAGATCCTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA
CAGTTGGATGAGTCTCAAGGTCCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCA
CTTCACGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAAC
CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCAGGGATTGTGCAACGCCA
GGCTTGTGCCAGCTTAATGGGAAC TGCTGTCTGGAACACACCACGGTGGAAAGTCAAGGCTG
CCCTGGAGGCTACTATGTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGG
TCATTTTATGACATCTGCGACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCAGTCAC
ATGCGCTCCAGGAAC TGCTAGGGCCTGACAGGCAGACATGCTTGATGAAAATGAATGTGA
GCAAAACAACGGTGGCTGAGTGAAGATCTGTGTGAACCTCAAAACTCCTACCGCTGTGAGTG
TGGGGTTGGCGTGTGCTAAGAAGTGAAGACTTGTGAAGACGTTGAAGGGATGCCACAA
TAACAATGGTGGCTGCAGCCACTCTTGCCTGGATCTGAGAAAGGCTACAGTGTGAATGTCC
CCGGGGCCTGGTGTCTGAGGATAACCACACTTCCAAGTCCCTGTGTTGTGCAAATCAA
TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTTGGCTGGAGCTTCTGACAGGCTACCCAA
CTCCTGCCAGGGAGTGTCCAACGGCACCCATGTCAACATCCTCTCTCAAGACATGTGG
TACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCAGCAACCTCGTACAGGCTACCCAA
GCAGACCCGGGAGCAGGGGACTTCATCCGAACCAGCAAGCTGCTGATCCGGTGAC
CTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACTGTTCCAAACCTTCGAAACCTCCC
ACTGGAAATCATGAGCCGAAATCATGGATCTTCCCATTCACTCTGGAGATCTCAAGGACAA
TGAGTTGAAGAGCCTACCGGGAGCTGTGAGCGGCTGGAAAGCTGGTGGAGAGCTGCTTGCAC
CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGATGGCTGTGTTCA
TGACTCGGTAAAGCAGTACACATCCGGGATCACCTAGCAAAGCACTTCCAGGTCCCTGTCTT
CAAGTTGTGGCAAAGACCACAAGGAAGTGTCTGACTGCCGGTTCTGTCTGGAGT
GTTGGACGAGCCTCCCGTGTGCCAGGGTTGCCACCGGCGAATGCGTGTGGGGAGGAGG
AGAGGACTCAGCCGGTCTACAGGGCAGACGCTAACAGGCGGCCGATCCGATCGACTGGGA
GGAC**TAG**TCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTTGGAGCTCTC
CCCCCACCAGCCCTCAAGAACATCTGCCAACAGCTGGTTCAAGACTCACACTGTGAGTTCA
ACTCCCAGCACCAACTCACTCTGATTCTGGCATCTAGTGGCAGGGCAGGTCAAGCACTGCTG
AACAAATGTGGCTGGGTGGGTTCTATCTTCTAGGGTTGAAAACACTAAACTGTCCACCCAGAA
AGACACTCACCCATTCCCTCATTCTTCTACACTAAATACCTCGTGTATGGTGAATC
AGACCACAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCCTAGAAAAATTAAACAG
TTACTGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAATTAGTTACCTT
GAAATTCAATTCAAATGCAGACTAATTAGGAATTGGAAAGTGTATCAATAACAGTAT
ATAATT

111/168

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCT
GCACTCGGGCCTCCTCCAGCCAGTGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACC
TGTGTGGGGAGGCCCTCCTGCTGCCTGGGGTACAATCTCAGCTCCAGGCTACAGGGAGACC
GGGAGGATCACAGAGCCAGC**ATG**TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG
ATGTCAAACCCCTGCGCAAACCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCA
TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCTCATCAAGGTGATTC
TGGATAAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCAGGAAGCAGCTGTGTG
ACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCAGTGTCAAGAGCTTCCCGAAG
GCCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA
CAGGGAACTGGTCTCTGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTA
GGCAGATGGCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGAAA
TCACAGAAAACAGCCAGGAGCTCGCATGCGGAACCTCAAGTGGCCCTGTCTCAGGCTCCC
TGGTCTCCCTGCACTGTCTGCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGGG
AGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTACGCATCCAGTACGACAAACAGCACGTCT
GTGGAGGGAGCATCCTGGACCCCCACTGGGCCTCACGGCAGCCCAGTGCCTCAGGAAACATA
CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAGCTTCCATCCCTGG
CTGTGGCCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACATCGCCC
TCATGAAGCTGCAGTCCCACTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTCT
TTGATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGA
ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCAGTCAGTCCAGGTATTGACAGCACACGGT
GCAATGCAGACGATGCGTACCAAGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCC
AAGGGGGTGTGGACACCTGCCAGGGTACAGTGGTGGCCCTGATGTACCAATCTGACCAAGT
GGCATGTGGTGGCATCGTAGCTGGCTATGGCTGCCGGCCAGCAGCCAGGAGTAT
ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG**TAAT**GCT
GCTGCCCTTGCACTGCTGGAGCCGCTTCCCTGCCCTGCCACCTGGGATCCCCCAA
AGTCAGACACAGAGCAAGAGTCCCCTGGTACACCCCTCTGCCACAGCCTCAGCATTTCT
GGAGCAGCAAAGGGCTCAATTCTGTAAGAGACCCCTCGCAGCCCAGAGGCAGGAGAG
TCAGCAGCCCTAGCTGCCACACTTGGTGTCCAGCATCCAGGGAGAGACACAGGCCACT
GAACAAGGTCTCAGGGTATTGCTAAGCCAAGAAGAACTTCCCACACTACTGAATGGAAGC
AGGTGTCTTGTAAAAGCCAGATCACTGTGGGTGGAGAGGAAGAGAAAGGGTCTGCCA
GCCCTGTCCGTCTTCAACCATCCCCAAGCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT
GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTATTGTTATTACAGCTATGG
CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA

FIGURE 113

GGCTGGACTGGAACTCCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGAT
TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTCAGCAACTAAAAAGCCACAG
GAGTTGAAC TGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT
AAAATCTGTTTTGTTCTTGTAACTAGCCTTACCTCCTAACACACAGAGGATCTGTCACT
GTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGT
CCCCTCGAAGCCGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAAC
TGTCTCACGTCTGGAGGCAGTGACTCGGGCAGTAGCAGGTAGCTGAGCCTCTGGTAGCTGCCG
CTTCAAGGTGGCCTGCCCTGGCGTAGAAGGGATTGACAAGCCCCAAGATTCAAGGCG
ATGGCTCCCAGTCCCCAGGCATCAGCCTGCTGTAGTCATCACTGCCCTGGGCCAGGACGG
GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTT
CATGTCCTGCACATCACCTGATCCATGGCTAATCTGAACCTGTCCCAAGGAACCCAGAGCT
TGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGA
CTTGCATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTG
TGCCAAATTATGGGTCAAGAAAAGATGGAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGC
ATTCA GTGGACATGTGGGGAAGGGCTGCCGATGGCGCATGACACACTCGGACTCACCTCTG
GGCCATCAGACAGCCGTTCCGCCGATCCACGTACAGCTGCTGAAGGGCAACTGCAGGC
CGATGCTCTCATGCCAGGCAGCAGCCAAATCTGCGATCACCAGCCAGGGCAGCCGTCTG
GGAAGGAGCAAGCAAAGTGACCATTCTCCTCCCTCCCTGTGAGAGGCCCTCTATGT
CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGCTAATGGCTCAGTGTGGCCAGGA
GGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCTGCTGTGCGAACACGGAAATGCCTCCAGT
AAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTAAATCATGTTCTAGT
AATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTCAAATGATCTCCAAGGGCCCT
TATACCCCAGGAGACTTGATTTGAATTGAAACCCCAAATCCAAACCTAAGAACCAAGGTGCA
TTAAGAACATCAGTTATTGCCGGGTGGTGGCTGTAATGCCAACATTGGGAGGCCAGGCG
GGTAGATCACCTGAGGTCAAGACCAAGGAGCTGGCTGTGCGAACATGGTGAACACGGAAATGCCTCCAGT
TACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCGCTGTATCCCAGCTACTCGGG
AGGCTGAGACAGGAGAATTACTGAAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT
TCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAATAAAAAAGAATTATGGTTATTT
GTAA

115/168

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGC
AAAGAACCTCCAGAGAATTGTGAAGACTGTCACATCTAAATGCAGAAGCTTTAAATCCAA
GAAAATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAAT
TGTCTGTTGGGGAGCAAGCAGTCTGCCGGAGGTACCCAAAAAGCCTATGACATGGA
GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAG
AACTGAAATATTCAAGCGGAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAAA
CGGATAACACTGGCATCTACTCGTGGGTCTCAAAATGTTTATCAAAACTCAGATTAAAGT
GATTCTGAATTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTT
CTTGAAACAGTCAGTGATTGGTCCCAGCAGAAAAGCCTATTGAAACCGAGATTCTTAA
AAATTCCAAAATTCTGGAGATTGTGATAACGTGACCATGTATTGGATCAATCCACTCTAAT
ATCAGTTCTGAGTTACAAGACTTGAGGAGGGAGAAGATCTTCACTTCCCTGCCAACGA
AAAAAAAGGGATTGAACAAAATGAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG
TCACGCCAGACAAGCAAGTGAGGAAGAACTCCAATAATGACTATACTGAAAATGGAATAGA
ATTGATCCCAGCTGGATGAGAGAGGTTATTGTTGATTTACTGCCGTGAGGCAACCGCTA
TTGCCGCCGTCTGTGAACCTTACTAGGCTACTACCCATATCCACTGCTACCAAGGAGG
ACGAGTCATCTGTCGTGTCATGCCTTGTAACTGGTGGTGGCCCGATGCTGGGAGGGT
CTAAAGGAGGTTGAGCTAAATGCTTAAACTGCTGGCAACATATAATAATGCATGCTATT
CAATGAATTCTGCCTATGAGGCATCTGCCCTGGTAGCCAGCTCTCCAGAATTACTGTAG
GTAATTCTCTTCAATAAAACTTACATTACCAAAAAAAAAAAAAAAA

117/168

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTCACACCTCGCAGCAGGAGGGCGGCAGCTCTCGC
AGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACATGCCAAGTGGTGGCGTTCTCC
TGTCCATCCTGGGCTGGCGGTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGG
ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA
GGCAGAGTTCAGGCTTCACCGAATGCAGGCCATTTCACCATCCTGGACTTCCAGCCATGC
TGCAGGCAGTGCAGGCCATTGATGATCGTAGGCATCGCCTGGTGCCATTGGCCTCTGGTAT
CCATCTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAGCCAACATGA
CACTGACCTCCGGGATCATGTTATTGTCTCAGGTCTTGCAATTGCTGGAGTGTCTGT
TTGCCAACATGCTGGTACTAACTCTGGATGTCCACAGCTAACATGTACACCGGCATGGTG
GGATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGCAGCTCTGGTGGCTGGTGG
CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACCAG
AAGAAACCAACTACAAAGCCGTTCTTATCATGCCCTCAGGCCACAGTGGCTACAAGCCTG
GAGGCTCAAGGCCAGCACTGGCTTGGTCCAACACCAAAAACAAGAAGATAACGATGGAG
GTGCCGCACAGAGGACGAGGTACAATCTTATGCCAACAGCACGACTATGT**TAA**TGCTCTA
AGACCTCTCAGCACGGCGGAAGAAACTCCCAGAGCTCACCCAAAAACAAGGAGATCCA
TCTAGATTCTTCTTGCTTTGACTCACAGCTGGAAAGTTAGAAAAGCCTCGATTTCATCTTG
GAGAGGCCAAATGGTCTTAGCCTCAGTCTGTCTAAATATTCCACCATAAAACAGCTGAG
TTATTATGAATTAGAGGCTAGCTCACATTCAATCCTCTATTCTTTAAATATAA
CTTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATCTATTCCAGCTTACCCAG
ACTCCCCCTCTCCTCCTAGTCAATAAACCCATTGATGATCTATTCCAGCTTACCCAG
AAAACTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTC
CCCACCCCCAACTGGCTAGTAATAAACACTACTGAAGAAGAAGCAATAAGAGAAAGATATT
TGTAATCTCTCCAGCCCATTGATCTGGTTCTACACTGTGATCTAAAGTACCAAACCA
AAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTGTTGACATCTCTTACAGC
AACACCATTCTAGGAGTTCCCTGAGCTCCACTGGAGTCCTTCTGCGGGTCAGAAA
TTGTCCTAGATGAATGAGAAAATTATTCTTAATTAAAGTCCTAAATATAGTAAATAA
ATAATGTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG
AAGGAAATGAAAAATAATTGCTTGACATTGTCTATATGGTACTTGTAAAGTCATGCTTAA
GTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTGGAGGCTGAGGAGGAAGG
ATCACTGAGCCCAGAAGTTGAGACTAGCCTGGCAACATGGAGAAGGCCGTGTCTACAAA
ATACAGAGAGAAAAATCAGCCAGTCATGGTGGCATAACACCTGTAGTCCAGCATTCCGGAG
GCTGAGGTGGAGGATCAATTGAGGCCAGGGAGGTTGGGCTGCAGTGAGCCATGATCACACC
ACTGCACCTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAGGAAATAATAATGGA
ACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTAA

119/168

FIGURE 119

GGAAAAACTGTTCTCTGTGGCACAGAGAACCGCTGCTCAAAGCAGAAGTAGCAGTTCCGG
AGTCCAGCTGGCTAAAACATCCCAGAGGATA**ATG**CAACCCATGCCTAGAAATCGCTGGG
CTGTTCTGGTGGTGGAAATGGTGGCACAGTGGCTGTCACTGTCACTGCCTCAGTGGAGA
GTGTCGGCCTCATGAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATG
AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTT
TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGCTGCTCCGTATGCCCTGGCT
TTCATGATGGCCATCCTGGCATGAAATGCACCAGGTGCACGGGGACAATGAGAAGGTGAAG
GCTCACATTCTGCTGACGGCTGGAATCATCTCATCACGGGCATGGTGGTGCATCCCT
GTGAGCTGGTTGCCAATGCCATCATCAGAGATTCTATAACTCAATAGTGAATGTTGCCAA
AAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCACTGGTGCATTGTTGGA
GGAGCTCTGTTGCTGCGTTTTGTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA
CCTTCCCATCGACAACCCAAAAAGTTATCACACCGAAAGAAGTCACCGAGCGTCACTCC
AGAAGTCAGTATGTG**TAG**TTGTATGTTTTAACTTACTATAAGCCATGCAAATGACA
AAAATCTATATTACTTCTCAAATGGACCCAAAGAAACTTGTATTACTGTTCTTAACACTGC
CTAATCTTAATTACAGGAACTGTGCATCAGCTATTATGATTCTATAAGCTATTCAGCAGAA
TGAGATATTAAACCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGG
TTCAAGCATCTACTCTTTTATCATTACTCAAATGACATTGCTAAAGACTGCATTATTT
ACTACTGTAATTCTCCACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCA
CATAGAGACATGCTTATGGTTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAAAT
AGAACTCAACTATTGCTTTCAGGGAAATCATGGTAGGGTTGAAGAAGGTTACTATTAATTG
TTTAAAACAGCTAGGGATTAATGTCCTCCATTATAATGAAGATTAAAATGAAGGCTTAA
TCAGCATTGTAAGGAAATTGAATGGCTTCTGATATGCTGTTTAGCCTAGGAGTTAGAA
ATCCTAACTTCTTATCCTCTCAGGGCTTGCATTCAAACACTGCTTCCAGGGCTACTC
AGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGTGTTAGGAAAGTGAAAATATTTT
GTTTTGTATTGAAGAAGAATGATGCATTGACAAGAAATCATATATGATGGATATATTT
TAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAATAAAAGAGCAGAAAATAT
GTCTGGTTTCATTGCTTACCAAAAAACAAACAACAAAAAGTTGTCTGTGAAAAATAAATT
ACCTGCTCCTATGTGGTACCTGAGTCAAAATTGTCACTTTGTTCTGTGAAAAATAAATT
CTTCTTGTACCATTCTGTTAGTTACTAAAATCTGTAAACTGTATTGTTCTGTGTTATT
CCAAATTGATGAAACTGACAATCCAATTGAAAGTTGTGCGACGTCTGTCTAGCTAAAT
GAATGTGTTCTATTGCTTATACATTATTAATAAATTGTACATTGTTCTAATT

121/168

FIGURE 121

GGAGAGAGGCGCGCGGGT GAAAGGC CATTGATGCAGCCTGC GGCGGCCTCGGAGCGCGGGCG
AGCCAGACGCTGACCACGTT CCTCCTCGGTCTCCGCCTCCAGCTCCGCCTGCCCGC
AGCCGGGAGCC**ATG** CGACCCCAGGGCCCCGCCCTCCCGCAGCGGCTCCGC GGCTCCTGC
TGCTCCTGCTGCTGCAGCTGCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAA
AGGCGCAGCTCCGGCAGAGGGAGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAG
CAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGGTACACCTGGGATCC
CAGGTGGGATGGATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCT
GGACACCCA ACTACAAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGGAAAA
TTGCGGAGTGTACATTACAAAGATGC GTCAAATAGT GCTCTAAGAGTTGTTCA GTGGCT
CACTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTCACTCAATGGAGCTG
AATGTTCA GGACCTCTCCCATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGA
ATTCAACAATTAATATTCA TCGCACTCTCTGTGGAAGGACTTGTGAAGGAATTGGT GCTG
GATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCTA
CTGGATGGAATTCA GTTCTCGCATCATTATTGAAGAACTACCAAAA**TAA** ATGCTTAATT
CATTTGCTACCTCTTTTATTATGCCTGGAATGGTCACTTAAATGACATTAAATAAG
TTTATGTATA CATCTGAATGAAAAGCAAAGCTAAATATGTTACAGACCAAGTGTGATTCA
CACTGTTTAAATCTAGCATTATTCA TTGCTCAATCAAAGTGGTTCAATATTTTT
TAGTTGGTTAGAATACTTTCTCATAGTCACATTCTCAACCTATAATTGGAATATTGTTG
TGGTCTTTGTTTTCTCTTAGTATAGCATTAAAAAAATATAAAAGCTACCAATCTTG
TACAATTGTAAATGTTAAGAATT TTTTATATCTGTTAAATAAAAATTATTCCAACA

123/168

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGGCTG
AACTGGGTGCTCATCACGGGAACGTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
CCAAATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTTAGGATCCAGTTTT
TTTTAACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAAGATGCAAAACGTAATATCCAT
GAAGATCCTATTACCTAGGAAGATTGATGTTGCTGCGAATGCGGTGTTGGATTATTT
GTTCTTGAGTGTCTGCGTGGCTGGCAAAGAATAATGTTCAAAATCGGTCCATCTCCAAG
GGTCCAATTTCCTGGTGTCAAGCAGCCCTGACTCACTACAGTGCAGCTGACAGGGG
CTGTATGCAACTGGCCCCCTAACGCCAAAGCACCTAACGGACGACCTTGAACAATACAA
AGG**ATG**GGTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCC
ACTGTCTACTGACAATGCTTCTTGCAGGATGCCCTAACGGCTGTAGGTGTGAA
GCCAAATGGTATATTGTAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT
TGCTTAGGTTGTCCTCGCTATAACACCCCTAACCTAAGTATAATCAATTAAAGGG
CTCAACCAAGCTCACCTGGCTATACCTGACCATAACCATATCAGCAATTGACGAAAATGCT
TTAATGGAATACGCAGACTCAAAGAGCTGATTCTAGTCCAATAGAATCTCTATTTCCT
AACAAATACCTTCAGACCTGTGACAAATTACGGAACCTGGATCTGTCCTATAATCAGTCAT
TCTCTGGGATCTGAACAGTTGGGGCTTGCAGGACTGCTGAGTTACATTACGGTCTAAC
TCCCTGAGAACCATCCCTGCGAATATTCCAAGACTGCCGAAACCTGGAACTTTGGACCTG
GGATATAACCGGATCCGAAGTTAGCCAGGAATGTCTTGCTGGCATGATCAGACTCAAAGAA
CTTCACCTGGAGCACAAATCAATTTCAGCTCAACCTGGCCCTTTCCAAGGTTGGCAGC
CTTCAGAACCTTACTTGCAGTGGAAATAAAACTAGTGTATAGGACAGACCAGTCCTGGACC
TGGAGCTCCTTACAAAGGCTGATTATCAGGCAATGAGATCGAAGCTTCAGTGGACCCAGT
GTTTCCAGTGTGCCCAGTCTGAGCGCCTAACCTGGATTCCAACAGCTCACATTATT
GGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAGTCTGCTGGGAATATATGG
GAATGCAGCAGAAATTGCTCCCTGTAAACTGGCTGAAAGTTAAAGGTCTAAGGGAG
AATACAATTATCTGTGCCAGTCCAAAGAGACTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
AACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTGATCTGGCCAGGGCTCTCCAAAG
CCGACGTTAACGCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTGGCCCCGACG
GTGGGAGGCCACAGAGCCGGCCCAGAGACCAGTGTGACGCCAGCACATCTTCCATAAA
ATCATCGCGGGCAGCGTGGCGCTTCTGTGCTCGTCACTGCTGGTTATCTACGTG
TCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAGCGCTCCCTCATGCGAAGGCAC
AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTATGAGAT
TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTAT
AACAAATCGGGCTCCAGGGAGTGTGAGGTAT**TGA**ACCATTGTGATAAAAGAGCTCTAAAGC
TGGGAAATAAGTGGTGTCTTATTGAACTCTGGTGACTATCAAGGAAACGCGATGCCCTC
CCCTCCCTCTCCCTCACTTGGTGGCAAGATCCTCCTGTCCGTTAGTGCATTATA
ATACTGGTCATTTCCTCTCATACATAATCAACCCATTGAAATTAAATACCACAATCAATGT
GAAGCTTGAACCTCGGTTAATATAACCTATTGTATAAGACCCTTACTGATTCCATTAAAT
GTCGCATTGTTAAGATAAAACTTCTTCATAGGTAAAAAAAAAA

125/168

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCGGAGGAGGGAGGGCTTTGCCGC
TGACCCAGAGATGGCCCCGAGCGAGCAAATCCTACTGTCCGGCTGCGCGCTACCGTGGCCG
AGCTAGCAACCTTCCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTC
TTGCTCGGTTGGAGACGGTGCAAGAGAATCTGCCCTATAGGGGAATGGTGCGCACAGCCC
TAGGGATCATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACA
GACACGTAGTGTATTCTGGAGGTGAATGGTCACATATGAACATCTCGAGAGGTTGTGTTG
GCAAAAGTGAAGATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTG
TATTGGCCAGTTTAGCCAATCCAAC TGACCTAGTGAAGGTTAGATGCAAATGGAAGGAA
AAAGGAAACTGGAAGGAAAACCATTGCGATTCGTGGTACATCATGCATTGCAAAATCT
TAGCTGAAGGAGGAATACGAGGGCTTGGCAGGCTGGTACCCAATATAACAAAGAGCAGCAC
TGGTGAATATGGGAGATTAAACCACTTATGATAACAGTGAAACACTACTTGGTATTGAATACAC
CACTTGAGGACAATATCATGACTCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTTCTA
TTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCACGAGATAAACAG
GAAGGGGACTTTGTATAAATCATCGACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGAT
TCATGAGTCTATATAAGGCTTTACCATCTTGGCTGAGAATGACCCCTGGTCAATGGTGT
TCTGGCTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCATTTAA

127/168

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCG
ACCCCCGTGTGGGCTAGTCCAGCGAGGCGACGGCGCGCTGGGCCATGCCAGGCCGGC
ATGGAGCGGTGGCGACCCGGCTGGCGCTGGTACGGGGCCTGGGGGCATCGGCGCGGC
GTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGACTGTGGCAAC
ATCGAGGAGCTGGCTGCTGAATGTAAGAGTCAGGCTACCCGGACTTGATCCCCTACAGA
TGTGACCTATCAAATGAAGAGGACATCCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC
GGTAGACATCTGCATCAACAATGCTGGCTGGCCGGCTGACACCCTGCTCTCAGGCAGC
ACCAGTGGTGGAAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGAA
GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGCACATATTAAACATCAATAGCATG
TCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTC
ACTGCGCTGACAGAGGGACTGAGGAAGAGCTCGGGAGGCCAGACCCACATCCGAGGCCACG
TGCATCTCTCCAGGTGTGGAGACACAATTGCCTCAAACACTCCACGACAAGGACCGCTGAG
AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCGAGGCTGTT
ATCTACGTCCCTCAGCACCCCCGACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAG
CAGGTGACCTAGTGACTGTGGAGCTCCTCCCTCCCCACCCCTCATGGCTGCCTCCTG
CCTCTGGATTTAGGTGTTGATTCTGGATCACGGGATACCACTTCTGTCCACACCCGACC
AGGGCTAGAAAATTGTTGAGATTTATATCATTTGTCAAATTGCTTCAGTTGAAATG
TGAAAAATGGCTGGGAAAGGAGGTGGTGTCCCTAATTGTTTACTTGTAACTTGTCTTG
TGCCCCCTGGCACTTGGCCTTGTCTCAGTGTCTCCCTTGACATGGAAAGGAGTT
GTGGCCAAAATCCCCATCTTGTGACCTCAACGTCTGGCTCAGGGCTGGGTGGCAGAGG
GAGGCCTTCACCTTATATCTGTGTTATCCAGGGCTCCAGACTTCCCTCTGCCTGCC
ACTGCACCCCTCTCCCCCTATCTATCTCCTCTGGCTCCCCAGCCCCAGTCTGGCTTCTGT
CCCCCTCTGGGTATCCCTCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCTGGC
CCAGTGGATTCATGGTGATCATTAAAAAGAAAAATCGCAACCAAAAAAAAAAAA

129/168

FIGURE 129

AACTTCTAC**ATG**GGCCTCCTGCTGCTGGTGTCTTCCTCAGCCTCCTGCCGGTGCCTACACC
ATCATGTCCTCCCACCCCTCCTTGACTGCGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGG
GAGCACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGCCTCGGCCAGAATTCCAGTTCTGGTT
TCATGCCAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTAAAGAGGGTT
TTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTCATTCAGTAGGCCACCAG
CCACCTGTGCCGTTGAGTGCTTGAAA**TGA**GGAACTGAGAAAATTAAATTCTCATGTATTTT
CTCATTATTTATTAAATTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGG
ATACATGTATAACAATATAATGATCAAATCAGGGTAACCTGGGATATCCATCACATCAAACAT
TTATTTTTATTCTTTAGACAGAGTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGGTGC
ATCTCAGCTTACTGCAACCTCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAA
GTAGCTGGGACTACAGGCATGCACCACAATGCCAACTAATTTGTATTTAGTAGAGACG
GGGTTTGCCATGTTGCCAGGCTGGCCTTGAACCTCCTGGCCTCAAACAATCCACTGCC
GCCTCCAAAGTGTATGATTACAGGCAGTGCAGCCACCGTGCCTGGCTAAACATTTATCTTT
CTTGTTGGAACTTGAAATTATAACATGAATTATTGTTAACTGTCATCTCCCTGCTGTG
CTATGGAACACTGGACTTCTCCCTCTATCTAACCTGTATATTGTACCAAGTAAACCAACCGT
ACTTCATCCCCACTCCTCTATCCTCCAAACCTCTGATCACCTCATTCTACTCTACCTC
CATGAGATCCACTTTTAGCTCCCACATGTGAGTAAGAAAATGCAATATTGTCTTCTGTG
CCTGGCTTATTCACTTAACATAATGACTCCTGTTCCATCCATGTTGCTGCAAATGACAGGA
TTTCGTTCTTAATTCAATTAAAATAACCACACATGGCAAAAA

131/168

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGATAAAGACCTAACACTGCTGACC**ATG**ATCAGCGCAG
CCTGGAGCATCTCCTCATCGGGACTAAAATTGGGCTGTTCCCTCAAGTAGCACCTCTATCAG
TTATGGCTAAATCCTGTCCATCTGTGTGCGATGCGGGTTCTTACTGTAATGATC
GCTTCTGACATCCATTCAAACAGGAATACCAAGAGGATGCTACAACCTCTACCTCAGAAC
ACCAAATAATAATGCTGGGATTCCCTCAGATTGAAAAACTGCTGAAAGTAGAAAGAATAT
ACCTATACCACAAACAGTTAGATGAATTCCCTACCAACCTCCAAAGTATGTAAGAGTTAC
ATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAAAAATTCCCTATCTGG
AAGAATTACATTAGATGACAACCTCTGTCTGCAGTTAGCATAGAAGAGGGAGCATTCCGAG
ACAGCAACTATCTCGACTGCTTCTGTCCCGTAATCACCTAGCACAATTCCCTGGGTT
TGCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTCATCACC
CTCTCAAGGTCTCACTAGTCTAAACGCCTGGTTAGATGGAAACCTGTTGAACAATCATG
GTTAGGTGACAAAGTTCTCAACCTAGTTAATTGACAGAGCTGCCCTGGCGGAATT
CCCTGACTGCTGCACCAGTAAACCTCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATA
ACCACATCAATCGGGTCCCCAAATGCTTTCTTATCTAAGGCAGCTCTATCGACTGGATA
TGTCCAATAATAACCTAAGTAATTACCTCAGGGTATCTTGTGATGATTGGACAATATAACAC
AACTGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT
TACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAAAGGTT
GTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGATTGAA
GCACCATTAGATAACCAC TGCAATACCCAAACACAGTGTATCCTGCCAAGGACAGTGGCCAG
CTCCAGTGCACAAACAGCCAGATATTAAGAACCCCCAAGCTCACTAAGGATCAACAAACCACAG
GGAGTCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATA
TCTCTGGAAACTTGCTTACCTATGACTGCTTGGACTCAGCTGGCTAAACTGGGCCATA
GCCCGGCATTGGATCTATAACAGAAACATTGTAACAGGGGACGGCAGTGAGTACTGGTCA
CAGCCCTGGAGCCTGATTCAACCTATAAAAGTATGCATGGTCCATGGAAACCAGCAACCTCT
ACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTCGAATGTACAACC
CTACAACCAACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCAAATTACCTTGG
CTGCCATATTGGTGGGCTGTGGCCCTGGTTACCATGCCCTCTGCTTGTGTTAGTGTGTTGGT
ATGTTCATAGGAATGGATCGCTTCTCAAGGAACGTGCATATAGCAAAGGGAGGAGAAAGAA
AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACACTATCGAAATCAGGGAAACTT
CTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCA
TATTTCTCCTAATGGAATGAATCTGTACAAAAAA CAATCACAGTGAAAGCAGTAGTAACCGAA
GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCA**TGA**TGCTGAAGGACTCACA
GCAGACTTGTGTTGGTTGGTTAAACCTAAGGGAGGTGATGGT

133/168

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCAGGCCACCCAGGCTCTGG
CAGCCCTGCCGGGCCACTTGTCTTC**ATG**TCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAG
GGCGTGAGAGGCAGTCTGGCTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGA
GCAGAGGCCGGCATGGCCAGCCTGGGCTGCTGCTCCTGCTCTTAAGTACAGCACTGCCACC
GCTGTGGCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT
GATCCTGTCTGCCTGGAGAGAGCCACCGTCTTAGAACAGAGGCTGCCTGAAATCACCT
GGATGGCATGGTGGGGTCCGAGTGTGAGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGC
CCAGGAGCCCCTGCTGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGC
TGCCATCCAGAGATCCCTCCACTACCTCAAGCTGAGTGTGATCCAAGTACCTAACAGAGAGTTCCA
GCTGACCCCTCCAGCCCAGGTTTGGAAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTT
GGTGTACCCACGTTGGGCCCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT
GCAGCTGCTGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGAGGAG
CCTCATGACCAAGCCCAGCTCAGGCTACTGCCTGTCCCACCAACTGCTCTTCTCCTCTG
GGCCAGAACATGAGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCACCTCTT
CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATGGATACGCCTACCCCTACCG
GGACATCTTATGGAAAACATCATGTTCTGGAATGGCGGTTCTCCGACTTCTAACAGCT
CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAACAGCAGGAAGGATGCTTCGGGGAGCCTGA
TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTCGAGGAGAGTGAA
GAGGCAGAAAAACAATTCCAGATTCTCGCTCTGGCTCAGGCTGGAGTACAGTGGCGCAA
TCTCGGCTCACTGCAACCTTGCCTCCTGGTTCAAGCAATTCTTGCCTCATCCTCCGAG
TAGCTGGACTACAGGAGCGTGCACCATACTGGCTAATTATTTATTTTTAGTAGAGAC
AGGGTTCATCATGTTGCTCATGCTGGCTCGAACCTCTGATCTCAAGAGATCCGCCACCTC
AGGCTCCAAAGTGTGGATT**TAG**GTGTGAGCCACCGTGTGGCTGAAAAGCACTTCAAA
GAGACTGTGTTGAATAAAGGGCCAAGGTTCTGCCACCCAGCACTCATGGGGCTCTCTCCCC
TAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGGTGGCTCTATACA
TCCTGGCAGAACACCCCCAGCAAACAGAGAGCCACACCCATCCACACGCCACCAAGCA
GCCGCTGAGACGGACGGTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTAGTCCTCA
TCCCTTAGATCTGGAGGGCACGGATCACATCCTGGAAAGAAGGCATCTGGAGGATAAGCAAA
GCCACCCGACACCCAAATCTTGGAAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGGGCCGGG
AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAACTGAAAAAA

135/168

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGCCGCTCTGTGGGCTTCTTCCGTCCTGCTGC
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCGGGGCTGCTGCTGAGGGATCGG
GAGGGAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGG
TGAAGCCTCAGGACTGGATCTGGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT
TCCTTAAGACAGATGGAGTTGTGGTCATGATATACTCTGGATCTTATGTAGTGGAAAG
TTGTATCTCCAGCTTACAGATTGATCCCCTCGAGTGGATATCACTCGAAAGGAAAAATGA
GAGCAAGATATGTGAATTACATCAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAA
TGAAATCTTCAGGTCCACCTTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTC
TAATGAACCCAATGGTTATGATGATGGTTCTCCTTATTGATATTGTGCTTCTGCCTAAAG
TGGTCAACACAAGTGATCCTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATT
CCAACCATGAGTTGCCTGATGTTCTGAGTTCATGACAAGACTCTCTCTTCAAAATCATCTG
GCAAATCTAGCAGCGCAGCAGTAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTCAG
GCCGTCCAGAGCTGGCATTGCACAAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAA
CCGTGTGAAGCAACTACTATAAACTTGAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTAGCACATGTTGTACTGGTACACGAGAAAACCCAGCTTCATCTTGCTGT
ATGAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTTATAGAAAATGCCATTAATAAAA
TTATATGAACTACTATACATTATGTATATTAAATTAAAACATCTTAATCCAGAAATCAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

137/168

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCGATTTCTCCCCAGTTCCCTGTGGGTCTGAGGG
GACCAGAAGGGT GAGCTACGTTGGCTTCTGGAAAGGGGAGGCTATATGCGTCAATTCCCCAAA
ACAAGTTTGACATTCCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTT
CCAGGCCTTACCTGCTGGCACTAACGGCGGAGCCAGGATGGGACAGAATAAAGGAGGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTCACGGGAG
GCTTGGCAGTTTCTTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGT
CTTGCCTCAGCCTTCTCTGCTGCGTTTATCTCCTATGGACTCCTCCACTGGACTGAAG
ACACTCAATTGGGAAGCTGTGTGATGCCACAAACCTTCAGGAAATACGAAATGGATTTC
GAGATACGGGGCAGTGTGCAAGCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCCATTGCTAAGACTC
TATCTGGACAGGGTATTAAAAACTACCAAGACCCCTGACCATTATACTCTCCGAAGATCAGC
AGCCTGCCAATTCTTCTTACCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACA
TGCCATTGTGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGGTTGGGGAACTAGACATTCTCTGCAATGGATGGAG
GAGACAGAATAGGAGGAAAGT GATGCTGCTGCTAAGAATATCGAGGTCAAGAGCTCCAGTCT
TCAATACTGCAGAGGAGGCATGACCCAAACCACCATCTCTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTGCATTACTGCTTGCATGATTGTCTTATGCATCCCC
AATCTTAATTGAGACCATACTGTATAAGATTTGTAATATCTTCTGCTATTGGATATATT
TATTAGTTAATATATTATTATTGCTATTAAATGTATTATTTTACTGGACATG
AAACTTAAAAAAATTCAAGATTATTTATAACCTGACTAGAGCAGGTGATGTATTTTAT
ACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGGCTAGGGGGTTATTCATTGTAT
TCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGATATTGAAATTGAACCAATGAC
TACTTAGGATGGGTGTGGAATAAGTTGATGTGGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTCTACAAATAAGTTCTGCATACCAAAAAAAAAAAAAAA

139/168

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAG
CCTAGCGTGTCCACG**ATG**CGGCTGGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAG
GTGCTAGGGATCGCGGTCTCCTCCTGGGATTCTTCCCGCTCCCGTTGCTCCTCTGCCAGA
GCGGAACACGGAGCGGAGCCCCAGCGCCGAACCCCTCGGCTGGAGCCAGTTCTAACTGGACC
ACGCTGCCACCACCTCTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGAT
TTTGTGTTGGGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTGTGGAAAAAGGA
GCATCTCACAGTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGC
TTGATGACGGGGAGCCTCTGGCTTGTCGACGTACAGGAACCTCAATTCTCCTGCACTG
CTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGAT
GAAACCTGGGTTAAATTATCCCACAGCATTGTGGAATATGATGGAACAACTCATTTC
GTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTGGATAAAAGTATTAAAAGA
GGAGATTGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGCCACATTTCAGGG
CCCAACAGCCCCCTGATTGGCAGAACAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCAC
ACCTCACTGCAGTCGAAGGAGAGAGAGACGCCCTTACCAATTGCTGGTCTTGTGGTGA
CATGGCATGTCGAAACAGGAAGTCACGGGCTCCTCCACCGAGGAGGTGAATACACCTCTG
ATTTAATCAGTTCTGCGTTGAAAGGAAACCCGGTGATATCCGACATCCAAAGCACGTCCAA
TAGACGGATGTGGCTGCGACACTGGCGATAGCACTGGCTTACCGATTCCAAAAGACAGTGT
GGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAGCAGTTGAGATTTC
TTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAGAAATGTGCCGTATATGAAAAGATCCT
GGGTTGAGCAGTTAAAATGTCAGAAAGATTGCAATGGAACTGGATCAGACTGTACTTGGAG
GAAAAGCATTCAAGTCCTATTCAACCTGGCTCCAAGGTTCTCAGGCAGTACCTGGATGCT
CTGAAGACGCTGAGCTGTCCCTGAGTGCACAAGTGGCCAGTTCTCACCCGTCTGCTCA
GCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCCAGTGTCACTCCTGGGTTT
CTCTGCTCTTATTGGTGTACCTGGTCTTCGGCGTTACGTCAATTGTGTGCACCTCAG
CTGAAAGTTCGTGTACTTCTGTGGCCTCTCGTGGCTGGCGAGGCTGCCTTGTGTTACCA
GAECTGGTTGAACACCTGGTGTGGCAAGTGCTGGCAGTGCCTGGACAGGGGGCCTCAGG
GAAGGACGTGGAGCAGCCTATCCAGGCCCTGGGTGTCCCACACAGGTGTCACATCTGT
GCTGTCAGGTCAAGATGCCTCAGTTGGAAAGCTAGGTTCTGCGACTGTTACCAAGGTGAT
TGTAAAGAGCTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAATCGGACA
GCCTCCCAGCAGAGGTGTGGAGCTGCAGTGAGGGAAAGAGACAATGGCCTGGACACTC
AGGAGGGTCAAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCAGAATGCATCCTGCC
TCATCAGGTCCAGATTCTTCAAGGCGACGTTCTGTGGAATTCTTAGTCCTGGCCT
CGGACACCTCATTGTTAGCTGGGAGTGGTGGTGGAGGCAGTGAAGAAGAGGGGATGGTCA
CACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCC
CCCACCCCAACCCCTGCACAGGCCCTCATCCCTCTGGCTTGAGCCGTAGAGGCCCTGTGCTG
AGTGTCTGACCGAGACACTCACAGCTTGTCATCAGGGCACAGGCTTCCCTGGAGCCAGGATG
ATCTGTGCCACGCTGACCTCGGGCCATCTGGGCTCATGCTCTCCTGCTATTGAATT
AGTACCTAGCTGCACACAGTATGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

141/168

FIGURE 141

GGCACGAGGCAAGCCTCCAGGTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTA
CAGAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTAC
AACTCCCTATAGAAAACAAC TGCCAGCACCTTAAGACCACACTCACACCTCAGAGTGAAGAAC
TTAAACCGAAGAAATT CAGCATT CATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGG
AATCTCATAGCAGTCCAGATAAAA ACTACATACGCCAGAGATCTTCTTG CATTAGCCTCA
TCCTTGAGCTCAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAG
TTTGTCTCTACTGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAG
AAACTGATGAAGCTGGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTTATAGGGCT
CAGGTGGGCTCCTGGAACATGCTGGAGTCGGCGGCTCACCCCGGATGGTCATCTGCACCTCC
TGCAATTGTAATGAGCCTGTTGGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATT
TCATTCAACCAGTTGCAAAGCTGAAATGAGCCCCAGTGAGGTAGCGATTAGGAAACTGCC
CCATTGAACGCCTCCTCGCTAATTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

143/168

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTT
AATCCAGGATCCTGTCCTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAG
TTTGTCCCACAGGGCTGTCTGAGCAGATAAGATTAAAGGGCTGGGTCTGTGCTCAATTAACTCC
TGTGGGCACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGGCC
TGCCGTGGAAGGGAGGTCTGCTGGCGCTGCTGCTGCTCTTAGGCTCCAGATCCTGC
TGATCTATGCCTGGCATTTCCACGAGCAAAGGGACTGTGATGAACACAATGTATGGCTCGTT
ACCTCCCTGCCACAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATG
CCTACAGACTGGGCACATCTTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCT
CAATGGAGCTACTGCTGGGAGAACTAGGTGTGGAAATTGAAAGACGACATTGACAACGCC
ATTTCCAAGAAAGCACAGAGCTGAACAATACTTACCTGCTTCTTACCATCAGCACCGAGC
CCTGGATGACTCAGTCAGCCTCTGAACAAAGACCTGCTGGAGGGATTCCACTGAGTGAAAC
CCACTCACAGGCTGTCCATGTGCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGA
GGACTCTCAGTGGCTGAGCAGCTTGGACTTGTATCCTATTGATGTGTTGAGA
TCTCAGATCAGTGTAGAAAATCCACACATCTTACGCTAATCATGTAGTAGATCATTAA
AACATCAGCATTAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

145/168

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACGCG**ATG**GAGGA
AGCACCTGAGCTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTGC
TCCAGACGAGGGGCATCAAGCACAGAACGAAAGTGGAACCGGAAGGCCCTGCCAGCAGCACTGCC
AGATCAGTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGC
TCGACATTGACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCT
GCATCAATGCCACCCAGGCCGAACCAGGGGAGTTCCAGAACGCCAGAACAGCTCCACC
AGCAGGTGCTCTGGCGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGTTGG
AGAGGGCGCAGGACTTCGGGTCAACATGCCAGCCAGTGCCTCTGCCTCTGGCTTGA
TCTGGCTCATGGTAAA**TAA**GCTTGCAGGAGGCTGGCAGTACAGAGCGCAGCGAGCAA
TCCTGGCAAGTGACCCAGCTCTCCCCAAACCCACGCCTGGTATGCGCTGCCT
GGCGATGCACTCGCACTGCAAATGCCGCTCCCACGTATGCCCTGGTATGCGCTGCCT
GATAGATGGGGACTGTGGCTCTCCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCA
CACTAGATTAGTAGTAAATGCTTGATGAGAACACATCAGGCACTGCCACCTGCTTCAC
AGTACTTCCAACAACACTCTTAGAGGTAGGTATTCCGTTTACAGATAAGGAAACTGAGGC
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGGCCAGGATTCAAC
CCTGGCTTGTCTAACCCAGGTTCTGCTCTGCTTAATTCCAGAGCTGTCTGGTATGACT
TATGTCTCACAGGGACCCACATCCAAACATGTATCTTAATGAAATTGTGAAAGCTCCATGTT
TAGAAATAATGAAAACACCTGA

147/168

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCTGGTCCAGAGTCTCAT
TTCCTGATGATTTAGACTCAAAGAAAACTC**ATG**TTCAGAAGCTCTTCTCTGGCCTC
CTCTCTGTCTTCTTCCCTTTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCA
AGCTGGAAATCTTCATTTGCTGTCACTGGGTAGGTCACTGAGTCTTAGTTTATT
TGAAATTCAACTTCAGATTCAAGGGTACATGTGAAGGTTGTTATGAGTATATTGCA**T**
GATGCTGAGGTTGGGT

149/168

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCCTACCTGGAGAC
TTGACTCCCGCGCCCCAACCTGCTTATCCCTGACCCTGAGTGTCAAGAGATCCTGCAGC
CGCCCAGTCCC GGCCCCTCTCCGCCAACACCCACCCCTCTGGCTTCTGTTTACTCC
TCCTTTCATTCATAACAAAAGCTACAGCTCCAGGAGGCCAGCGCCGGCTGTGACCCAAGCC
GAGCGTGGAAAGA**ATG**GGGTTCTCGGGACCGGCACCTGGATTCTGGTGTAGTGCTCCGATT
CAAGCTTCCCCAACCTGGAGGAAGCCAAGACAATCTACATAATAGAGAATTAAGTGCA
GAAAGACCTTGAATGAACAGATTGCTGAAGCAGAAGACAAGATTAAAAAACATATCCT
CCAGAAAACAAGCCAGGTCAAGAGCAACTATTCTTGTGATAACTGAACTGCTAAAGGCA
ATAACAGAAAAGGAAAAATTGAGAAAGACAATCTATAAGAAGCTCCCCACTTGATAAT
AAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACGTGATGATTATGACTCT
ACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCATCAACTAGACGGG
ACTCCTTAACCGCTGAAGACATTGTCCATAAAATCGTGCAGGATTATGAAGAAAATGAC
AGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGCCAA
GCACATACACTGGAAGATGAAGTAGCAGAGGTTACAAAATTAACTCAAGGAAGCCAAAC
AATTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCA
GAGAAAGTGACTIONGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAACGATGAAACA
GTATCTAACACATTAACCTTGACAAATGGCTGGAAAGGAGAACTAAACCTACAGTGAAGAC
AACTTGAGGAACCTCAATATTCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAA
AAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAACACTGATTGACTTGTGAAG
ATGATGGTGAATATGGAACAATATCTCCAGAAGAAGGTGTTCTACCTTGAAAACCTGGAT
GAAATGATTGCTCTCAGACCAAAAACAAGCTAGAAAAAATGCTACTGACAATATAAGCAAG
CTTTCCCAGCACCCTCAGAGAAGAGTCATGAAGAAACAGACAGTACCAAGGAAGCAGCT
AAGATGGAAAAGGAATATGGAAGCTTGAGGATTCCACAAAAGATGATAACTCCAACCCAGGA
GGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTGGAAGCCATCAGAAAAAATT
GAATGGTTGAAGAAACATGACAAAAAGGAAATAAGAAGATTATGACCTTCAAAGATGAGA
GACTTCATCAATAACAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAAGAAGCC
GAGGCCATCAAGCGCATTAGCAGCCTG**TAAA**ATGGCAAAAGATCCAGGAGTCTTCAAC
TGTTTCAGAAAACATAATATAGCTTAAACACTCTAATTCTGTGATTAAAATTGGACCC
AAGGGTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTTACAAGTGGTAAAACATAGC
TTTCTTCCCGTAAAAACTATCTGAAAGTAAAGTGTATGTAAGCTGAAAAA
AAA

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCCTGTGGAGCTC
AAGATGGTCTGAGTGGGGCGCTTGCTTCGAATGAAGGACTCGGCATTGAAGGTGCTTTAT
CTGCATAATAACCAGCTTAGCTGGAGGGCTGCATGCAGGGAAAGTCATTAAAGGTGAAGAG
ATCAGCGTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCAG
GGTGGAAAGCCAGTGCCTGTCATGTGGGGTGGGCAGGAGCCACTCTAACACTAGAGCCAGTG
AACATCATGGAGCTCTATCTGGTGCAAGGAATCCAAGAGCTTCACCTCACCGGGGGAC
ATGGGGCTCACCTCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCGTGCACGGTGCCT
GAAGCCGATCAGCCTGTCAGACTCACCCAGCTCCCGAGAATGGTGGCTGGAATGCCCATC
ACAGACTCTACTTCCAGCAGTGTGACTAGGGCAACGTGCCCGAGAACTCCCTGGGAGAG
CCAGCTCGGGTGGGGGTGAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAG
GACCCCCACGTCTGACTTAGTGGGCACCTGACCACTTGTCTCTGGTCCCAGTTGGATAA
ATTCTGAGATTTGGAGCTCACTCCACCGCTCTCCCCACTGGATGGTGCCTACTGCTGTGGAAC
CTTGAAAAACCATGTGGGTAAACTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGTG
GGGGAGTGGTGGGAATCATTCTGCTTAATGGTAAGTACAAGTGTACCCCTGAGCCCCCAG
GCCAACCCATCCCCAGTTGAGCCTATAGGGTCAGTAGCTCCACATGAAGTCCTGTCACTC
ACCACTGTGCAGGAGAGGGAGGTGGTCATAGAGTCAGGGATCTATGCCCTGGCCAGCCCC
ACCCCTCCCTTAATCCTGCCACTGTCAATGCTACCTTCCTATCTCTCCCTCATC
TTGTTGTGGGCATGAGGAGGTGGTGTAGTCAGAAGAAATGGCTCGAGCTCAGAAGATAAAAGA
TAAGTAGGGTATGCTGATCCTCTTTAAAAACCAAGATAACATCAAATCCCAGATGCTGGT
CTCTATCCCCATGAAAAGTGCTCATGACATATTGAGAAGACCTACTTACAAAGTGGCATATA
TTGCAATTATTTAATTAAAAGATAACCTATTATATTTCTTATAGAAAAAGTCGGAA
GAGTTTACTTCAATTGTAGCAATGTCAGGGTGGCAGTATAGGTGATTTCTTTAATT
TGTAAATTATCTGTATTCCTAATTCTACAATGAAGATGAATTCTTGATAAAATAA
GAAAAGAAATTAACTTGAGGTAAGCAGAGCAGACATCATCTCTGATTGTCCTCAGCCTCCAC
TTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTCTGGTTGGTTGTAGTAGTGT
GGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTGTGGCTGGAATCT
CTGGTAAGGAACTTAAAGAACAAAATCATCTGGTAATTCTTCTAGAAGGATCACAGCCC
CTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGAATTGTGTC
CCCTCAAATTCACATCCTCTTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAGGTCTCTG
CAGATGTAGTTAGTTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATATGACTG
GTTCTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGGGGAAAGACTATGAAAG
ATGAAGGCAGAGATCGGAGTTTGAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACCA
TCAGAAGCTTGGAAAGAGGCAAAGAAGAATTCTCCCTAGAGGGCTTAGAGGGATAACGGCTCT
GCTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGATAAATTGGCTGTT
TAAGCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATC
CCTGTCTCCTCGTGTTCACATTCTGTGTGTCCTCCACAATGTACCAAAGTTGTCTTTG
TGACCAATAGAATATGGCAGAAGTGTAGGGATGCCACTTCAAGATTAGGTATAAAAGACAC
TGCAGCTTCTACTTGAGCCCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACT
CGCTCTGGGGAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAA
AATGAAGTCTCCTGCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATC
GATGTTGTTGTTAAGTTGCTCAGTTGGTCTAACTTGTATGCAGCAATAGATAAATAA
TATGCAGAGAAAGAG

153/168

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCC
GCCCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCCTGCCACCAGCTGCCTCCTCTC
TTGGCCCTCTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCACTGCAGGCTTGACAAG
TCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTG
GCTGATAACAAACACAGACGTTCGTCTCATGGGGAGAAACTGTTCCACGGAGTCAGTATGAGT
GAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAA
TCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCTTCTGGCCAGGCTCAGAACAGG
CTAACGACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAG
GACACAGTAAAAAGCTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTG
TTTATGTCTCTGAGAAATGCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCC
CCTTCCCTGCTAGAAATAACAATTAGATGCCCAAAGCGATTTTTAACAAAAGGAAGA
TGGGAAGCCAAACTCCATCATGATGGGTGGATTCAAATGAACCCCTGCGTTAGTTACAAAGG
AAACCAATGCCACTTTGTTATAAGACCAGAAGGTAGACTTCTAACGATAGATATTATTG
ATAAACATTTCATTGTAAGGGTTCTATACACAGAAAACAATTATTAAATAATTGTC
TTTTCCATAAAAAAGATTACTTCCATTCTTAGGGAAAAACCCCTAAATAGCTTCATG
TTCCATAATCAGTACTTATATTATAATGTATTATTATTATAAGACTGCATTTAT
TTATATCATTATTAAATATGGATTATTATAGAAACATCATTGATATTGCTACTTGAGTG
TAAGGCTAATATTGATATTATGACAATAATTAGAGCTATAACATGTTATTGACCTCAA
TAAACACTGGATATCCC

155/168

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCT
GTCAGTCAGTCCCCACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTCACTGCAGA
GGGCTGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGG
CCAAGCTGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAAACTGGATCCCAGGGGGAGG
GTGCAGATGAGGGAGCGACCCAGATTAGGTGAGGAACAGTTCTCTCATTAGCCTTCTACAG
GTGGTTGCATTCTGGCAATGGCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGC
CCCAGCAAAGGGCAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCAGTGTGCCTGTGCCTCCC
CTAGAGCCTGCTAGGCCAACGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTC
AACAGCAGGCCATCTCCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAG
GACCTGTACCACGCCGTTGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCACATG
GACCCCCGGGCAACTCGGAGCTGCTCTACCACAACCAGACTGTCTTCTACAGGCAGGCCATGC
CATGGCGAGAAGGGCACCAAGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTA
GCTTGTGTGTGTGCGGCCCCGTGTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTT
TTGGGAAACCTGGAGCCAGGTGTACAACCACCTGCCATGAAGGGCAGGATGCCAGATGCTT
GGCCCTGTGAAGTGCTGTGGAGCAGCAGGATCCGGGACAGGATGGGGGCTTGGGAA
AACCTGCACTTCTGCACATTTGAAAAGAGCAGCTGCTGCTTAGGGCCGGAAAGCTGGTGT
CCTGTCATTTCTCTCAGGAAAGGTTCAAAGTTCTGCCATTCTGGAGGCCACCACCTCCT
GTCTCTTCCCTTTCCATCCCTGCTACCCCTGGCCAGCACAGGCACTTCTAGATATTC
CCCCCTGCTGGAGAAGAAAGAGCCCTGGTTTATTTGTTACTCATCACTCAGTGAGC
ATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTTGTACATGGATGATTCTGAGGAGGA
AGCTGTTATTGAATGTATAGAGATTATCAAATAATCTTATTAAAAATGAAAAAA

157/168

FIGURE 157

CCGGCGATGTCGCTCGTGTGCTAACGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAG
CCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA
ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGAC
TATTCAATTTGATGAATGTAAGCTGGGTACTCCGGCAGATGCCAGCATCCGCTTGTGAAG
GCCACCAAGATTGTGTGACGGGAAAAGCAACTCCAGTCCTACAGCTGTGAGGTGCAAT
TACACAGAGGCCTCCAGACTCAGACCAGACCCCTGGTGGTAAATGGACATTTCTACATC
GGCTTCCCTGTAGAGCTGAACACAGTCTATTCTATTGGGCCATAATATTCTAATGCAAAT
ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCTCACAGGCTGCCTAGACCACATA
ATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAACCTGTGGGATCCGAACATCACTGCTTGT
AAGAAGAATGAGGAGACAGTAGAAGTGAACCTCACACCACCTCCCTGGGAAACAGATACTG
GCTCTTATCCAACACAGCACTATCATGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAA
CAAACGCGAGCTTCAGTGGTATTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTG
ACTCCATATTTCTACTTGTGGCAGCGACTGCATCCGACATAAGGAACAGTTGTGCTCTGC
CCACAAACAGCGTCCCTTCCCTCTGGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCT
CTCCTCCTGCTGTCTGCTGGGCCACATGGTGCTGGTGGCAGGGATCTATCTAATGTGG
AGGCACGAAAGGATCAAGAAGACTCCTTTCTACCAACCACACTGCCCCCATTAAAGGTT
CTTGTGGTTACCCATCTGAAATATGTTCCATCACACAATTGTTACTTCACTGAATTCTT
CAAAACCATTGCAGAAGTGAGGTACCTTGAAAGTGGCAGAAAAAGAAAATAGCAGAGATG
GGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCAGCAGACAAAGTCGTCTCCTCTTCC
AATGACGTCAACAGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAAC
TCTCAAGACCTCTCCCCCTGCCTTAACCTTTCTGCAGTGATCTAAGAAGCCAGATTCT
CTGCACAAATACGTGGTGGCTACTTAGAGAGATTGATACAAAGACGATTACAATGCTCTC
AGTGTCTGCCCAAGTACCCACCTCATGAAGGATGCCACTGCTTCTGTGCAGAACTTCTCCAT
GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCTGCTCCTTG
TAG

159/168

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTG
CTGTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGGAAAATCCCCAAAGTAGGA
CATACTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGAC
ATTGGCATCATCAATGAAAACCAGCGCGTTCCATGTCACGTAACATCGAGAGCCGCTCCACC
TCCCCCTGGAATTACACTGTCACTTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCC
CAGTGTAGGAACCTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTT
CCCATCCAGCAAGAGACCCCTGGTCTGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAG
TTGGAGAAGGTGCTGGTGACTIONTGGCTGCACCTGCGTACCCCTGTCATCCACCATGTGCAG
TAAGAGGTGCATATCCACTCAGCTGAAGAAG

161/168

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCA
GGACAGAGAGTGACAAAACATACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGG
ATTCCAGCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCTGGGGGGGG
CAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGAT**ATG**CCTGTGCCCTGGTTCTT
GCTGTCCTTGGCACTGGGCCAAGCCCAGTGGCTTCTCTGGAGAGGCTGTGGGCCCTCA
GGACGCTACCCACTGCTCTCCGGGCTCTCCTGCCCTCTGGGACAGTGACATACTCTGCCT
GCCTGGGGACATCGTGCCTGCTCCGGGCCCTGCTGGCGCTACGCACCTGCAGACAGAGCT
GGT GCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC
CGTGCATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGG
GGTGGAGGAGCCTAGGAATGCCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCTACCC
TACTGCCGCTGCGTCTGCTGGAGGTGCAAGTGCCTGCTGCCCTGTGCAAGTTGGTCAGTC
TCTGCCCTCTGTGCTATATCACTCCTCCACCCCTGCCCTACCCACTCACCTACCAATCTGGTC
CTATACTCAGCCCAGGTACGAGAAGGAACCTAACACACAGCAGTCGCTGCCCTGCCCTG
GCTCAACGTGTCAGCAGATGGTACAACGTGATCTGGTTCTGAATGTCTCTGAGGAGCAGCA
CTTCGGCCTCTCCCTGTAATCAGGTCAGGGCCCCAAAACCCGGTGGCACAAAAAA
CCTGACTGGACCGCAGATCATTACCTTGAAACCACACAGACCTGGTCCCTGCCTCTGTATTCA
GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCTTCAGGGAGGACCCCG
CGCACACCAGAACCTCTGGCAAGCCGCCCCACTGCGACTGCTGACCCCTGCAGAGCTGGCTGCT
GGACGCACCGTGCCTGCGCTGCCCGAGAAGCAGGACTGTGCTGGCGGCTCCGGGTGGGACCC
CTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACGTACTGTGGACAAGGTTCTCGAGTT
CCCATTGCTGAAAGGCCACCCCTAACCTCTGTGTTAGGTGAAACAGCTCGGAGAAGCTGCAGCT
GCAGGAGTGTGTTGGGCTGACTCCCTGGGCCCTCTCAAAGACGATGTGCTACTGTTGGAGAC
ACGAGGGCCCCCAGGACAACAGATCCCTCTGTGCTGGACCCAGTGGCTGTACTTCACCTACC
CAGCAAAGCCTCACGAGGGCAGCTCGCCTGGAGAGTACTTACTACAAGACACTGCAGTCAGG
CCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTGCCCATGGACAAATA
CATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTGCCGCTGCGCTTCCCT
CATCCTCCTCTCAAAAGGATCACCGAAAGGGTGGCTGAGGCTCTGAAACAGGACGTCCG
CTCGGGGGCGGCCAGGGCGCGCGGCTCTGCTCCTCTACTCAGCCGATGACTCGGTTT
CGAGCGCCTGGTGGCGCCCTGGCGTCGGCCCTGTGCCAGCTGCCGTGCGCGTGGCGTAGA
CCTGTGGAGCGCTGTAAGTGGAGCGCTATGGGCTGCCCTGCCTACTCTTGCCGCTGCGCTGTG
CCAGACCCCTGCAGGAGGGCGCTGGTGGCTTGTCTCTCTCCGGTGCAGTGGCGCTGTG
CGCCTCGCTCAGCTGCGTGTGCCGACTTCTTGCAAGGGCCGGCGCCGGCAGCTACGTGGG
GGCCTGCTCGACAGGCTGCTCCACCCGGACGCCGTACCCGCCCTTTCCGACCGTGCCCGT
CTTCACACTGCCCTCCCAACTGCCAGACTTCTGGGGCCCTGCAGCAGCCTCGCGCCCGCG
TTCCGGGGCGCTCCAAGAGAGAGCGGGAGCAAGTGTCCCGGGCCCTCAGCCAGCCCTGGATAG
CTACTTCCATCCCCGGGACTCCCGCGCCGGACGCCGGTGGGACCAGGGCGGGACCTGG
GGCGGGGGACGGGACT**TAA**ATAAAGGCAGACGCTTTTCTAAAAAAA

163/168

FIGURE 163

GGGAGGGCTCTGCCAGCCCCG**ATG**AGGGACGCTGCTGACCATCTGACTGTGGGATCCCTGG
CTGCTCACGCCCTGAGGACCCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACT
TTGAAAACATCCTGACGTGGACAGCGGGCAGAGGGCACCCAGACACGGTCTACAGCATCG
AGTATAAGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTAGCGGATCACCGGA
AGTCCTGCAACCTGACGGTGGAGACGGCAACCTACGGAGCTACTATGCCAGGGTACCGCT
GTCAGTGCAGGGAGGGCCGGTCAGCCACCAAGATGACTGACAGGTTAGCTCTGCAGCACACT
ACCCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTGAGATGATTGTCAT
CCTACCCCCACGCCAATCCGTGCAGGCATGGCCACCGGCTAACCCCTGGAAGACATCTCCAT
GACCTGTTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAAG
CAGAGAGAATATGAGTTCTCGGCCTGACCCCTGACACAGAGTTCCCTGGCACCACATGATT
TGCAGTCCACCTGGCCAAGGAGAGTGCCTTACATGTGCCAGTGAAAGACACTGCCAGAC
CGGACATGGACCTACTCCTTCTCGGAGCCTTCTGTCCATGGGCTTCCTCGCAGTA
CTCTGCTACCTGAGCTACAGATATGTCACCAAGCCGCTGCACCTCCAACTCCCTGAACGTC
CAGCGAGTCCTGACTTCCAGCCGCTGCCTCATCCAGGAGCACGTCTGATCCCTGTCTT
GACCTCAGCGGCCAGCAGTCTGGCCAGCCTGTCAGTACTCCAGATCAGGTGTCTGGA
CCCAGGGAGCCGCAGGGAGCTCCACAGCAGCAGCAGTCCATGTCAGTGACACTTGGCAG
CCAGACATCTCCATCCTCAGCCCTCAAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCC
TATGCCCAAACGCTGCCCTGAGGTGGGCCACAGGCACTCTAAGGTCCAGCCTCCATGCCC
GCTCAATTCCCATTCTACGCCCCACAGGCCATCTCAAGGTCCAGCCTCCATGCCC
CAAGCCACTCCGGACAGCTGGCCTCCATGGGTATGCATGGAAGGTTCTGGCAAAGAC
TCCCCCACTGGGACACTTCTAGCTAAACACCTAGGCTAAAGGTCAAGCTCAGAAAGAG
CCACCAGCTGGAAGCTGCATGTTAGGTGGCTTCTGTCAGGAGGTGACCTCCTGGCTATG
GAGGAATCCCAAGAAGAAAATCATTGCACCCAGGCTGGGATTTGCACAGACAGAACATCT
GACCCAAATGTGCTACACAGTGGGAGGAAGGGACACCACAGTACCTAAAGGGCCAGCTCCCC
CTCCTCTCCTCAGTCCAGATCGAGGGCCACCCATGTCCTCCCTTGCAACCTCCTCCGGT
CCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGCTGCTGGAGTCCCTGTGTGTC
AAGGATGAAGCCAAGAGCCCAGCCCTGAGACCTCAGACCTGGAGCAGCCCACAGAACATGGAT
TCTCTTCTAGAGGCCTGGCCCTGACTGTGAGTGGGAGTC**TGA**GGGAATGGAAAGGCTT
GGTGCCTCCTCCCTGTCCTACCCAGTGTCACTCCTGGCTGTCAATCCCAGTCCTGCCAT
GCCACACACTCTCGCGATCTGGCCTCAGACGGGTGCCCTGAGAGAAGCAGAGGGAGTGGCATG
CAGGGCCCTGCCATGGTGCCTCCTCACCGAACAAAGCAGCATGATAAGGACTGCAGCGG
GGGAGCTCTGGGAGCAGCTGTAGACAAGCGCGTGCTCGCTGAGCCCTGCAAGGAGAAA
TGACAGTGCAAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGGCCACCTCTAACACCA
TGGATTCAAAGTGCTCAGGGAATTGCTCTCCTTGCCCCATTCTGGCCAGTTACAATCT
AGCTCGACAGAGCATGAGGCCCTGCCCTGTCATTGTCATTGTCAGCTCATTCTCTGATAGAACAAAG
GGAAAGAACCGGCCCTGGAAAAGAACCGAGAAGGAGGCTGGCAGAACAGAACACTGCACT
TCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGAGCTCAT
TCCCGCCAGGGCAACTGCCTGACGTTGCACGATTGTCAGCTCATTCTCTGATAGAACAAAG
CGAAATGCAGGTCCACCAGGGAGGGAGAACACACAAGCCTTCTGCAGGCAGGAGTTGAGAC
CCTATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAA
CACACTGTACTGATGTCACAACATTGCAAGCTCTGCCCTGGGTTGAGCCATCTGGCTCAA
TTCCAGCCTCACCACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTC
GGTTGCTCATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAA
ATGAAGTCATGCTTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCCAATAACGGT
AGCTATTTAAAAAAAAAAAAAA

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAGTCACCCGGGCCGCGGTGCCACAAC**AT**
GGCTGCGCGCCGGGCTGCTCTGGCTGTTCGTGCCTGGGGCGCTCTGGTGGTCCCAGG
CCAGTCGGATCTCAGCACGGACGGCGTTCTCGGACCTCAAAGTGTGCAGGGACGAAGAGTG
CAGCATGTTAATGTACCGTGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTGT
GAATTAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAAC
TTGGGCTGGAAGTGTGAACACAGTTGGATATTCACAAAGATTTGATCAAGGTACTTCA
TAAATACACGGAAGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAGG
'AGGAAGAGATGATTTAATAGTTATAATGTAGAACAGCTTTAGGATCTTGGAACTGGAGGA
CTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAACAGTTCTCAGCACAGAGAGAAATCTCCTGA
GGAGTCTCGGGGCGTGAACCTGACCCTGTGCCTGAGCCGAGGCATTAGAGCTGATTAGA
GGATGGAGAAGGTGCTTCTCAGAGAGCACCGAGGGCTGCAGGGACAGCCCTAGCTCAGGA
GAGCCACCTCACACCAAGCGGTCTGCCTAAGCAGCTCAGGGAGTGCAGTCAGTCTGGACAC
TTTGAAAGAAATTCTGCACGATAAAATTGAAAGTGCCTGGAAAGCGAAAGCAGAACTGGCAATAG
TTCTCCTGCCTCGGTGGAGCGGGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAG
TCAGAGAGGAAGTGGACAGTGCCTTATTCAAGCAAAGGATTGCTGGCATCAAATCT
AAGTTGTTACAAAGATTGTTT**TAG**ACTAAGCTGCCTGGCAGTTGCATTGGCAGGAA
AAACAAAAATATATTATTCACAAAGGATTGCTGGCAGTTGCATTGGCAGGAA

167/168

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGC
GCAACGGTGTGGTCCAAGCCGGGCTCTGCCTCGCCTCTAGGACATACACGGGACCCCTAA
CTTCAGTCCCCAAACGCGCACCTCGAAGTCTTGAACCTCCAGCCCCCACATCCACGCGCG
CACAGGCGCGCAGGGCAGGTCCCAGGCGAAGGGCATGCGCGAGGGGGTCGGGCAGCTGG
GCTCGGGCGGCGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGG
CTGGCCCTGGCAGAGGCCCTCGCTCACGCAACACCTGCTGCTGCCACCGCGCCCGCA**A**
TGAGCCGCGTGGTCTCGCTGCTGGCGCCGCGCTGCTGCGGCCACGGAGCCTTCTGCC
GCCGCGTGGTCAGCGGAAAAGGTGTGTTTGCTGACTCAAGCATCCCTGCTACAAAATGG
CCTACTCCATGAACGTCCAGCCAGTGAAGCTTCAGGAGGCACGCCTGGCTGTGAGAGTG
AGGGAGGAGTCCTCCTCAGCCTGAGAATGAAGCAGAACAGAAAGTTAATAGAGAGCATGTTGC
AAAACCTGACAAAACCGGGACAGGGATTCTGATGGTGTGATTCAGGATAGGGCTTGGAGGA
ATGGAGATGGGAAACATCTGGTGCCTGCCAGATCTCTACCAAGTGGCTGATGGAAGCAATT
CCCAGTACCGAAACTGGTACACAGATGAACCTTCTCGGAAGTGAAGAAGTGTGTTGTGATGT
ATCACCAACCAACTGCCAACCTGGCCTTGGGCTCCCTACCTTACCAAGTGGATGAAACAGCCCTG
GGTGTAAACATGAAGCACAATTATATTGCAAGTATGAACCAGAGATTAATCCAACAGCCCTG
TAGAAAAGCCTTACCAACACCAGGAGACACCCATCAGAATGTGGTTACTGAAG
CAGGTATAATTCCAATCTAATTATGTTGTTACCAACAATACCCCTGCTCTACTGATAC
TGGTTGCTTTGGAACCTGTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAAAAACTA
GTCCAACCCAGTCTACACTGTGGATTCAAAGAGTACCAAGAAAAGTGGCATGGAAGT**A**
AATAACTCATTGACTTGGTCCAGAATTGTAATTCTGGATCTGTATAAGGAATGGCATCAG
AACAAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGCAAGATGAACGTGTAAGCTCCCCCT
TGAGGCAAATATTAAAGTAAATTGTTATATGCTATTATTCATTTAAAGAATATGCTGTGCTA
ATAATGGAGTGAAGACATGCTTATTGCTAAAGGATGCAACCCAAACTCAAAACTCAAGCAA
TGAAATGGACAATGCAGATAAAGTGTATCAACACGTCGGGAGTATGTGTGTTAGAAGCAAT
TCCTTTATTCTTCACCTTCATAAGTTGTTATCTAGTCATGTAATGTATATTGATTGA
AATTTCAGTGTGAAAGTATTTCACCTTGCTAAAGTGTGATAAAAGTAAAGCAACTGTTCTA
ATATTATTATTATGGCATCTCATTTCAATAACATGCTTTGATTAAAGAAACTTATTAC
TGTGTCAACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTGTTCTCGAA
ATAATTCATCTTCAGCTCTGCTTTGGTCAATGCTAGGAAATCTTCAGAAATAAGA
AGCTATTCTATTAAAGTGTGATATAAACCTCCTCAAACATTACTTAGAGGCAAGGATTGTCT
AATTCAATTGTGCAAGACATGTGCCCTATAATTATTAGCTTAAACAGATTGTT
TAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCACTGCAATTGAAACAAAGT
GACATACACAATATAATCATATGCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTG
AGGGTTCTGAAATCAATGTGGTCCCTCTCTGCCACTAAACAAAGATGGTTGTTGGGGTTT
GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGC
CTCTGACTATTAGTATAAAAGAGGTCAATGTTGAGACCAGGTGAATAGTCACATCAG
TGTGGAGACAAGCACAGCACAGACATTAGGAAGGAAAGGAACATACGAAATCGTGTGAAA
ATGGGTTGGAACCCATCAGTGTGATCGCATATTGATGAGGGTTGCTTGAGATAGAAAATG
GTGGCTCCTTCTGTCTTATCTCCTAGTTCTCAATGCTTACGCCTGTTCTCAAGAGA
AAGTTGTAACCTCTGGTCTTCAATGTCCTGTGCTCCTTTAACCAAATAAGAGTTCTTG
TTTCTGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16318 A3

(51) International Patent Classification ⁷ :	C12N 15/12. C07K 14/47, 14/705, G01N 33/53, C12N 15/62, C07K 16/18	PCT/US00/04414	22 February 2000 (22.02.2000)	US
		PCT/US00/05601	1 March 2000 (01.03.2000)	US
		60/187,202	3 March 2000 (03.03.2000)	US
		60/191,007	21 March 2000 (21.03.2000)	US
		PCT/US00/08439	30 March 2000 (30.03.2000)	US
		60/199,397	25 April 2000 (25.04.2000)	US
		PCT/US00/14042	22 May 2000 (22.05.2000)	US
		60/209,832	5 June 2000 (05.06.2000)	US
(21) International Application Number:	PCT/US00/23328			
(22) International Filing Date:	24 August 2000 (24.08.2000)			
(25) Filing Language:	English			
(26) Publication Language:	English			
(30) Priority Data:				
	PCT/US99/20111			
	PCT/US99/21090	1 September 1999 (01.09.1999)	US	
	60/169,495	15 September 1999 (15.09.1999)	US	
	60/170,262	7 December 1999 (07.12.1999)	US	
	60/175,481	9 December 1999 (09.12.1999)	US	
	PCT/US00/04341	11 January 2000 (11.01.2000)	US	
	PCT/US00/04342	18 February 2000 (18.02.2000)	US	
		18 February 2000 (18.02.2000)	US	
(71) Applicant (for all designated States except US):	GENEN-TECH, INC. [US/US] , 1 DNA Way, South San Francisco, CA 94080-4990 (US)			
(72) Inventors; and				
(75) Inventors/Applicants (for US only):	EATON, Dan, L. [US/US] ; 75 Knight Drive, San Rafael, CA 94901 (US). FILVAROFF, Ellen [US/US] ; 538 18th Avenue, San Francisco, CA 94121 (US). GERRITSEN, Mary, E. [CA/US] ; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US] ; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US] ; 2627 Easton Drive, Burlingame, CA 94010 (US). GRIMALDI, Christopher, J. [US/US] ; 1434 36th			

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

<subunit 1 of 1, 266 aa, 1 stop
<MW: 29766, pI: 8.39, NX(S/T): 0
MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIA
AVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAV
LTFGMGSLYMFVQTILSYQMQPPIHGKQVFWRLLLVIWCGVSALSMLTCSSVLHSGNFGTDL
EQKLHWNPEDKGYVLHMITTAAEWSMSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPC
PINNERTRLLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites:

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 01/16318 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/23328

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C07K14/705 G01N33/53 C12N15/62
C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 25825 A (BOUGUELERET LYDIE ;GENSET SA (FR); DUCLERT AYMERIC (FR); DUMAS MIL) 27 May 1999 (1999-05-27) the whole document ---	1-20
X	WO 99 24836 A (ENDRESS GREGORY A ;HUMAN GENOME SCIENCES INC (US); FENG PING (US);) 20 May 1999 (1999-05-20) the whole document ---	1-20
A	EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document ---	
A	WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document ---	
		-/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

24 January 2001

23.04.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/23328

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 35-38, in as far as they pertain to in vivo methods, are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-20 (all partially).

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: Invention 1: 1-20, all partially

PRO180: nucleic acid with seq.ID.1, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.2 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.2 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide.

2. Claims: Inventions 2-76: claims 1-20, all partially

Subject matter as defined for invention 1, but related to the respective nucleic acid/polypeptide sequences of:

2. PRO218, represented by seq.ID.s 3 and 4,
3. PRO263, represented by seq.ID.s 5 and 6,
4. PRO295, as represented by seq.ID's 7 and 8,
5. PRO874, as represented by seq.ID's 9 and 10,
6. PRO300, as represented by seq.ID's 11 and 12,
7. PRO1864, as represented by seq.ID's 13 and 14,
8. PRO1282, as represented by seq.ID's 15 and 16,
9. PRO1063, as represented by seq.ID's 17 and 18,
10. PRO1773, as represented by seq.ID's 19 and 20,
11. PRO1013, as represented by seq.ID's 21 and 22,
12. PRO0937, as represented by seq.ID's 23 and 24,
13. PRO842, as represented by seq.ID's 25 and 26,
14. PRO1180, as represented by seq.ID's 27 and 28,
15. PRO831, as represented by seq.ID's 29 and 30,
16. PRO1115, as represented by seq.ID's 31 and 32,
17. PRO1277, as represented by seq.ID's 33 and 34,
18. PRO1074, as represented by seq.ID's 35 and 36,
19. PRO1344, as represented by seq.ID's 37 and 38,
20. PRO1136, as represented by seq.ID's 39 and 40,
21. PRO1109, as represented by seq.ID's 41 and 42,
22. PRO1003, as represented by seq.ID's 43 and 44,
23. PRO1138, as represented by seq.ID's 45 and 46,
24. PRO994, as represented by seq.ID's 47 and 48,
25. PRO1069, as represented by seq.ID's 49 and 50,
26. PRO1411, as represented by seq.ID's 51 and 52,
27. PRO1129, as represented by seq.ID's 53 and 54,
28. PRO1027, as represented by seq.ID's 55 and 56,
29. PRO1106, as represented by seq.ID's 57 and 58,
30. PRO1291, as represented by seq.ID's 59 and 60,
31. PRO3573, as represented by seq.ID's 61 and 62,
32. PRO3566, as represented by seq.ID's 63 and 64,
33. PRO1098, as represented by seq.ID's 65 and 66,
34. PRO1158, as represented by seq.ID's 67 and 68,
35. PRO1124, as represented by seq.ID's 69 and 70,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

represented in seq.ID.156 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.156 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO10272 using its interaction with PRO5801 (seq.ID.158), method for linking a bioactive molecule to a cell expressing PRO10272 through the use of PRO5801, and method of modulating at least one activity of said cell thereby.

5. Claims: Invention 78: claims 1-3,5-12,14-38, all partially

PRO20110: nucleic acid with seq.ID.159, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.160 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.160 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO20110 using its interaction with PRO20040 (seq.ID.162), method for linking a bioactive molecule to a cell expressing PRO20110 through the use of PRO20040, and method of modulating at least one activity of said cell thereby.

6. Claims: Invention 79: claims 1-3,5-12,14-38, all partially

PRO10096: nucleic acid with seq.ID.153, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.154 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.154 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO10096 using its interaction with PRO20233 (seq.ID.164), method for linking a bioactive molecule to a cell expressing PRO10096 through the use of PRO20233, and method of modulating at least one activity of said cell thereby.

7. Claims: Invention 80: claims 1-3,5-12,14-38, all partially

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: Invention 83: claims 1-3,5-12,14-38, all partially

PRO20233: nucleic acid with seq.ID.163, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.164 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.164 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO20233 using its interaction with PRO10096 (seq.ID.154), method for linking a bioactive molecule to a cell expressing PRO20233 through the use of PRO10096, and method of modulating at least one activity of said cell thereby.

11. Claims: Invention 84: claims 1-3,5-12,14-38, all partially

PRO1890: nucleic acid with seq.ID.167, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.168 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.168 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO1890 using its interaction with PRO19679 (seq.ID.166), method for linking a bioactive molecule to a cell expressing PRO1890 through the use of PRO19679, and method of modulating at least one activity of said cell thereby.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/23328

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0073454 A		AU 2883600 A	28-09-2000
		WO 0053756 A	14-09-2000
		WO 0078961 A	28-12-2000
		WO 0104311 A	18-01-2001
		AU 1749800 A	04-10-2000
		AU 1749900 A	12-07-2000
		AU 2399300 A	28-09-2000
		AU 3107000 A	19-06-2000
		WO 0105972 A	25-01-2001
		WO 0032778 A	08-06-2000
		WO 0055319 A	21-09-2000
		WO 0037638 A	29-06-2000
		WO 0105836 A	25-01-2001
		WO 0053751 A	14-09-2000
		WO 0077037 A	21-12-2000
		WO 0109327 A	08-02-2001
		WO 0075317 A	14-12-2000
		AU 5922999 A	03-04-2000
		WO 0118210 A	15-03-2001
		WO 0116318 A	08-03-2001
		WO 0119987 A	22-03-2001
		AU 6498499 A	03-04-2000
		WO 0119991 A	22-03-2001
		AU 2224800 A	28-09-2000
		AU 2596700 A	28-09-2000
		AU 3107700 A	28-09-2000
		WO 0053754 A	14-09-2000
		WO 0053750 A	14-09-2000
		AU 2474700 A	19-06-2000